DOCTOR OF PHILOSOPHY

Mechanisms of repression of the transcription factor NRF2 by KEAP1- and B-TrCP-dependent ubiquitin ligases and how the dysregulation of NRF2 contributes to lung cancer progression

Robertson, Holly

Award date:
2019

Mechanisms of repression of the transcription factor NRF2 by KEAP1- and β-TrCP-dependent ubiquitin ligases and how the dysregulation of NRF2 contributes to lung cancer progression

Holly Robertson BSc (Hons) Biomedical Sciences

Thesis submitted for the degree of Doctor of Philosophy

University of Dundee

January 2019

Table of Contents

List of figures

List of tables

Abbreviations

1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP)

2-cyano-3,12-dioxooleana-1,9-dien-28-oic acid (CDDO)

2-mercaptoethanol (2ME)

3-(4,5-dimethylthiazol-2-Yl)-2,5-diphenyltetrazolium bromide (MTT)

5,5'-dithio-bis-2-nitrobenzoic acid (DNTB)

8-oxo-deoxyguanosine (8-oxo-dG)

Activating transcription factor 6 (ATF6)

Activator protein-1 (AP-1)

Additionally NRF2-complexed hypomorph (ANCHOR)

Adenocarcinoma (ADC)

Adenomatous polyposis coli (APC)

Adenosine diphosphate (ADP)

Adenosine triphosphate (ATP)

Aldo-keto reductase (AKR, AKR1B10, AKR1C1, AKR1C2 and AKR1C3)

Alzheimer’s disease (AD)

Ammonium persulfate (APS)

AMP-activated protein kinase (AMPK)

Amyotrophic lateral sclerosis (ALS)

Antioxidant response element (ARE)

Apoptosis regulating kinase -1 (ASK-1)

Apoptotic protease activating factor 1 (Apaf-1)

Aryl-hydrocarbon receptor (AhR)

B-cell lymphoma 2 (BCL-2)

Basic-region leucine zipper (bZIP)

Beta-transducin repeat containing protein (β-TrCP, β-TrCP1 and β-TrCP2)

Bithio-bis-2-nitrobenzoic acid (DNTB)

Bovine serum albumin (BSA)

Broad complex, tram-track, bric-a-brac (BTB)

BTB and CNC homology (BACH, BACH1 and BACH2)

Buthionine sulfoximine (BSO)

Butylated hydroxyanisole (BHA)

c-Jun N-terminal kinase (JNK, JNK1, JNK2 and JNK3)

C-terminal region (CTR)

Cancer cell line encyclopaedia (CCLE)

Cancer research UK (CRUK)

Cap'n'collar (CNC)

Cap'n'collar basic-region leucine zipper (CNC-bZIP)

Cap'n'collar protein (CncC)

Carbon dioxide (CO_2)

Casein kinase (CK, CK1 and CK2)

Catalase (CAT)

CDDO-Imidazolide (CDDO-Im)

CDDO-Methyl Ester (CDDO-ME)

CDK-inhibitors (CKIs)

Chromodomain helicase DNA binding protein 6 (CHD6)

Chronic Granulomatous disease (GCD)

Clustered regularly interspaced short palindromic repeats (CRISPR)

Computed tomography (CT)

Copper (Cu)

Copper-zinc superoxide dismutase (CuZnSOD)

CREB binding protein (CBP)

Cullin 1 (CUL1)

Cullin 3 (CUL3)

Cyclin-dependant kinases (CDKs and CDK9)

Cyclohexamide (CHX)

Cytoplasmic (C)

Delta Ct (dCt)

Detergent compatible (DC)

Diabetes mellitus (DM)

Diacylglyerol (DAG)

Dimethyl fumerate (DMF)

Dipeptidyl peptidase 3 (DPP3)

Dual specificity tyrosine-phosphorylation-regulated kinase (DYRK, DYRK1A, DYRK2, DYRK3 and DYRK4)

Dulbecco's modified eagle medium (DMEM)

Dulbecco’s phosphate-buffered saline (DPBS)

Endoplasmic reticulum (ER)

Endoplasmic reticulum associated degradation (ERAD)

Enhanced chemiluminesence (ECL)

Erythroid cell-derived protein with CNC homology (ECH)

Escherichia coli (E.Coli)

Eukaryotic initiation factor 2B (eIF2B)

Extracellular SOD (ECSOD)

F-Box and WD repeat domain-containing 7 (FBW7)

F2-Isoprostanes (F2-isoPs)

Falvin adenine dinucleotide ($FADH_2$)

Flavin adenine dinucleotide (FAD)

Foetal calf serum (FCS)

Frozen robust multiarray analysis (fRMA)

Gain-of-function (GOF)

Glucose-6-phosphate dehydrogenase (G6PD)

Glucose-6-phosphate sodium salt (G-6-P)

Glutamate-cysteine ligase (GCL)

Glutamate-cysteine ligase catalytic (GCLC)

Glutamate-cysteine ligase modifier (GCLM)

Glutaminase (GLS and GLS2)

Glutathione peroxidase (GPX, GPX2)

Glutathione reductase (GR)

Glutathione S-transferase alpha 2 (GSTA2)

Glutathione S-transferase (GST)

Glutathione synthase (GS)

Glycogen synthase kinase (GSK-3)

Hanks balanced salt solution (HBSS)

Homologues to the E6AP carboxyl terminus (HECT)

Horseradish peroxidase (HRP)

Human NRF2 (hNRF2)

Hutchinson-Gifford progeria syndrome (HGPS)

Hydrogen peroxide (H_2O_2)

Hydrogen sulphide (H_2S)

Hydroxyl radical (HO·)

Hypoxia inducible factor-1 (HIF-1)

Hypoxia inducible factor-1 alpha (HIF-1α)

Hypoxia inducible factor-1 beta (HIF-1β)

Intervening region (IVR)

Kelch-like ECH-associated protein 1 (KEAP1)

Kirsten rat sarcoma (KRAS)

Knock out (KO)

Leucine rich repeats (LRR)

Luria Broth (LB)

Manganese (Mn)

Manganese superoxide dismutase (MnSOD)

Mechanistic target of rapamycin (mTOR)

Mediator subunit 16 (MED16)

Metalloproteinases (MMPs, MMP2 and MMP9)

Mitochondrial targeting sequence (MTS)

Mitogen-activated protein kinase (MAPK)

Mixed-linage kinase-1 (MLK1)

Monochloromobimane (mBC)

Mouse NRF2 (mNRF2)

Multidrug resistance-associated protein 1 (MRP1)

Murine double minute 2 (MDM2)

Musculoaponeurotic fibrosarcoma (Maf)

Myelin Basic Protein (MBP)

N-acetylcysteine (NAC)

N-terminal domain (NTD)

N-terminal region (NTR)

N, N, N',N'-Tetramethylethylenediamine (TEMED)

NAD(P)H dehydrogenase quinone 1 (NQO1)

Nicotinamide adenine dinucleotide phosphate (NADP)

Nicotinamide adenine dinucleotide phosphate (NADPH)

Nicotinamide adenine dinucleotide phosphate oxidase (NOX)

Non-small cell lung cancer (NSCLC)

Not significant (ns)

NRF2-ECH homology (Neh)

Nuclear (N)

Nuclear factor erythroid 2-related factor 1 (NRF1 and *NFE2L1*)

Nuclear factor erythroid 2-related factor 2 (NRF2 and *NFE2L2)*

Nuclear factor erythroid 2-related factor 3 (NRF3 and *NFE2L3*)

Nuclear factor-erythroid 2 p45 subunit (NF-E2 p45)

Nuclear factor-kappa beta (NF-κβ)

Nuclear Magnetic Resonance (NMR)

Okadaic acid (OA)

Optical density (OD)

Oxidized glutathione (GSSG)

Oxygen (O_2)

Papillary renal cell carcinoma (PRCC)

Parkinson’s disease (PD)

Pentose phosphate pathway (PPP)

Perocymonocarbonate (HCO_4^-)

Peroxyl radical (ROO•)

Phosphate buffer saline (PBS)

Phosphogluconate dehydrogenase (PGD)

Polymerase chain reaction (PCR)

Polyvinyl difluroride (PVDF)

Protein kinase C (PKC)

Protein phosphatase 2A (PP2A)

Reactive oxygen species (ROS)

Real time PCR (RT-PCR)

Really interesting new gene (RING)

Reduced glutathione (GSH)

Retinoid x receptor alpha (RxRα)

Ribosomal protein S6 kinase (S6K)

RING box protein 1 (RBX1)

RING-between RING (RBR)

RING-finger protein 4 (RNF4)

S-phase-kinase-associated protein-1 (SKP1)

Sestrin 1 (SESN1)

Sestrin 2 (SESN2)

Seven in absentia homolog 2 (SIAH2)

Short interfering RNA (siRNA)

Sirutin (SIRT1)

Skinhead-1 (SKN-1)

Skp, cullin, F-box (SCF)

Small cell lung cancer (SCLC)

Small Maf (sMaf)

Small ubiquitin-related modifier (SUMO1 and SUMO2)

Sodium dodecyl sulfate (SDS)-polyacrylamide gel (SDS-PAGE)

Sodium fluoride (NaF)

Sodium vanadate (Na_2VO_3)

Solute carrier family 3 member 2 (*SLC3A2*)

Solute carrier family 7 member 11 (xCT, encoded by the gene *SLC7A11*)

Squamous cell carcinoma (SQCC)

Standard deviation (StDev)

Standard error of the mean (SEM)

Stress-activated protein kinase (SAPK, SAPK2A, SAPK2B, SAPK3 and SAPK4)

Sulfhydryl group (SH)

Sulforaphane (SFN)

Superoxide anion radical ($O2^{\cdot-}$)

Superoxide dismutase (SOD, SOD1, SOD2 and SOD3)

TANK-binding kinase 1 (TBK1)

Taurosodeoxycholic acid (TUDCA)

Tert-butylhydroquinone (t-BHQ)

The Cancer Genome Atlas (TCGA)

Thioredoxin (Trx, Trx1 and Trx2)

Thioredoxin reductase (TrxR, TrxR1 and TrxR2)

Thr/tyr kinase (TTK)

TPA-response element (TRE)

Tumour necrosis factor alpha (TNF-a)

Type 1 DM (T1DM)

Type 2 DM (T2DM)

Tyrosine kinase inhibitors (TKIs)

Ubiquitin-activating enzymes (E1's)

Ubiquitin-conjugating enzymes (E2's)

Ubiquitin-proteasome system (UPS)

Ubiquitin ligases (E3's)

Ultraviolet (UV)

Unfolded protein response (UPR)

Vascular endothelia growth factor (VEGF)

Vitamin E radical (Vit E-O•)

Xenobiotic response element (XRE)

Zinc (Zn and Zn^{2+})

β-D-1-thiogalactopyranoside (IPTG)

γ-glutamyltranspeptidase (GGT)

Acknowledgements

I would like to start by thanking my supervisor Professor John Hayes for giving me the opportunity to carry out my PhD in his lab and also for his constant support. I hope to take forward the strong scientific grounding that you have given me into my next job and will always be grateful for your guidance. I would also like to thank Professor Albena Dinkova-Kostova for her scientific contributions to the development of the projects in this PhD and for her kindness.

The first project in this thesis, on the role of a priming kinase in NRF2 degradation, would not have been possible without the great mind that is Doctor Calum Sutherland. Thank you Calum, for all the hours you spent teaching me the radioactive experiments, your vast amount of patience and also for your guidance and support. Doctor James Hastie also contributed greatly to this work by providing all the kinases.

I was extremely luckily that my PhD was part-funded by AstraZeneca and I was able to carry out a three-month internship at their Hodgkin site at Chesterford. I would like to thank my industrial supervisor Doctor Stephen Durant for all of his support throughout the PhD and during my placement at AZ. Also, Doctor Robert McEwan who helped with the bioinformatics analysis and Doctor Molly Taylor who helped with the Fluidigm experiments (not shown in this thesis).

The generation of the isogenic panel of CRISPR/Cas9 cell lines would have not been possible without Doctor Laureano de la Vega. Thank you for providing some of the constructs, for teaching me the protocol and also for your patience. Thank you to Mr Thomas McCartney for carrying out the sequencing of the CRISPR/Cas9 cell lines and being so helpful in the optimisation of the genomic DNA extraction protocol and the analysis of the results. Also thank you to Ms Maureen Higgins (aka the NQO1 assay master) for teaching me the NQO1 assay and Miss Claudia Bento-Pereira for teaching me the Cycloheximide protocol and keeping up the morale in the lab.

I would like to thank all past and present members of the Hayes lab. Especially Doctor Ritu Sharma, who always made time to help me and in times that I thought every experiment was just awful pointed me back on the right path.

I have had the absolute delight of being a founding member of the “Little fun level 6 office” and would like to thank my dear friends: Morven, Hugh, Aparajitha, Dianne, Maureen and Amber. It has been a right laugh sitting in that office and I will miss you more than I can say! Also thank you to: Angus, Daniella, Kevin and Sarah for providing top quality banter and their friendship.

Finally, I would like to thank my friends and family for their overwhelming support and love. My best friends Megan and Rachel who have had to put up with all my PhD chat for years, you guys are the best! Jason my extremely patient, understanding and caring boyfriend, thank you for putting up with all my aggy-ness and mad working hours. I am super excited for our next adventure together! I am really lucky to have such a mad, caring, supportive and loving family. Thank you to my mum and dad who have had to put up with so much ranting down the phone over the past four years. You have both always supported me in whatever I wanted to do, and I really couldn’t have done it without you. Thank you to my brother, Ben, who always makes me laugh with his awful jokes and definitely keeps me very grounded. I miss you lots!

The Phd has been a rollercoaster at times but I will always look back fondly on the experience, thank you.

Declaration

I declare that this thesis is based on results obtained from investigations which I have personally carried out in the Cancer division of the School of Medicine in the University of Dundee from October 2014 to September 2018 using funding provided by the Biotechnology and Biological sciences research council (BBSRC) and AstraZeneca. I declare that the entire thesis is my own composition. Any work other than my own is clearly stated in the text and acknowledged with reference to any relevant investigators or contributors. This thesis has never been presented previously, in whole or in part, for any award of any higher degree. I have consulted all the references cited within the text of this thesis.

Signed …….. Date ……..

I confirm that Holly Robertson has spent the equivalent of at least 9 terms in the Division of Cancer Research, School of medicine, University of Dundee, and that she has fulfilled the conditions of the University of Dundee, thereby qualifying her to submit this thesis in application for the degree of Doctor of Philosophy.

Signed……... Date ……

Abstract

Lung cancer is the leading cause of cancer related mortality worldwide and since the discovery of the important role that NRF2 plays in regulating the expression of phase II drug metabolism genes, NRF2 has been a highly studied therapeutic target. NRF2 signalling is often elevated in lung cancer due to mutations in *NFE2L2,* the gene that encodes NRF2; and *KEAP1*, the cytoplasmic regulator of NRF2. Since it became established that elevated NRF2 signalling provides a beneficial role to the tumour through enhancing proliferation, enabling survival in highly oxidative conditions and enhancing resistance to chemotherapeutic drugs; many groups have focussed on understanding KEAP1-independent forms of NRF2 regulation.

It is now widely accepted that GSK-3 phosphorylation of the Neh6 domain of NRF2 is fundamental for $SCF^{\beta\text{-}TrCP}$-mediated degradation of the transcription factor. Previous work by other groups has indicated a potential role for a priming kinase to pre-phosphorylate the Neh6 domain and subsequently enhance GSK-3 mediated phosphorylation and resulting 26S proteasomal degradation. In the first results chapter of this thesis it was demonstrated using the DYRK family of isoenzymes that the presence of a priming kinase enhances GSK-3 phosphorylation and stimulates NRF2 degradation. Also, it has been shown that DYRK family members carry out priming of NRF2 through phosphorylating Ser-347, which enhances subsequent phosphorylation of Ser-342 and Ser-338 by GSK-3; regulating both NRF2 activity and stability. Additionally, it was shown that altering the phosphorylation status of Ser-347 impacts cell proliferation and chemo-sensitivity to platinum-based compounds. The ability of other kinases to phosphorylate the Neh6 domain was also revealed.

In the second results chapter of this thesis a bioinformatics style approach was utilized to analyse the impact of mutations in *KEAP1, NFE2L2* and *CUL3*; in terms of their effect on NRF2 activity using the expression of NRF2-target genes as a read out. The following points were highlighted from this analysis; (1) *KEAP1* mutations are more prevalent than *NFE2L2* and *CUL3* mutations in both lung cancer cell lines and tumours; (2) *KEAP1* and *NFE2L2* mutations vary in there associated zygosity's, with *KEAP1* mutations being predominantly homozygous and *NFE2L2* mutations being predominantly heterozygous; (3) mutations in

KEAP1 coexist with mutations in *KRAS* whereas, mutations in *NFE2L2* co-exist with mutations in *TP53* and (4) lung cancer cell lines and tumours harbouring mutations in *KEAP1* have greater expression of NRF2-target genes than those harbouring either *NFE2L2* or *CUL3* mutations.

In the third and final results chapter, both a panel of commercial cell lines and an in-house generated CRISPR/Cas9 panel of cell lines have been shown in multiple experiments to validate the bioinformatics findings of the previous chapter. Also, in this chapter it was identified that *KEAP1* mutant cells are more reliant on the glutathione biosynthesis pathway for survival and therefore more sensitive to alterations in this pathway.

The work demonstrated in this thesis has revealed new insights into the complex regulation of NRF2 and the importance of studying the mutations that lead to aberrant NRF2-signaling in lung cancer. Through showing the importance of a priming kinase in the regulation of KEAP1-independent degradation of NRF2, a completely novel alternative way of stimulating NRF2 degradation in the presence of a *KEAP1* mutation is possible. The data displayed in the second and third results chapters highlights the importance of understanding the impact that mutations in *KEAP1* and *NFE2L2* have on NRF2 signalling; and suggests that theses mutations should not be grouped together in terms of the effect. Additionally, the data presented in the second and third results chapters highlights the importance of generating physiologically relevant NRF2 models. This greater understanding of the regulatory mechanisms in place to control the levels of NRF2 in the cell and the mutations that lead to aberrant NRF2-signalling in cancer will hopefully aid in the development of new therapeutic strategy to combat this disease.

1.0 Introduction

1.1 Oxidative stress

All aerobic life needs to produce energy to survive. Aerobic metabolism is often utilized as it produces a greater energetic yield of 18 times more adenosine triphosphate (ATP) from a single molecule of glucose, than that of anaerobic metabolism. However, this larger energetic yield comes at a cost to the cell with the production of potentially damaging reactive oxygen species (ROS) such as: hydrogen peroxide (H_2O_2), hydroxyl radical ($HO^{\bullet}$), superoxide anion radical ($O_2^{\bullet-}$) and singlet oxygen (O_2), as bi-products of the incomplete oxidation of oxygen (Halliwell, 2007). The $O_2^{\bullet-}$ anion is formed through the one electron reduction of O_2 and is the precursor for the formation of other forms of ROS. Superoxide can also combine with nitric oxide, which is endogenously produced in the cell, to form extremely damaging peroxynitrite (Squadrito and Pryor, 1995). H_2O_2 is formed either spontaneously or through dismutation of superoxide, catalysed by the enzyme superoxide dismutase (SOD). Also, highly reactive singlet oxygen has been shown to react with ascorbate leading to the formation of H_2O_2 (Kramarenko et al., 2006). H_2O_2 is considerably less reactive than other forms of ROS but therefore is more stable and has the potential to travel in to the cell nucleus and damage DNA (Halliwell, 2007;Turrens, 2003). Also, in the presence of carbon dioxide (CO_2) H_2O_2 forms perocymonocarbonate (HCO_4^-), another potentially damaging radical (Bakhmutova-Albert et al., 2010). Incomplete reduction of H_2O_2 leads to the formation of highly reactive and unstable $HO^{\bullet}$. In an attempt to increase stability, $HO^{\bullet}$ rapidly scavenges electrons from neighbouring molecules, resulting in a damaging chain reaction of electron scavenging.

A homeostatic balance usually exists within the cell between the production of ROS and the antioxidant defences. Oxidative stress occurs when this balance is disrupted and the levels of endogenous ROS over powers the antioxidant defence capacity of cells, leading to damage of to several cellular components such as: DNA, protein and lipids, which can ultimately result in the development of disease.

1.1.1 Sources of reactive oxygen species

The majority of endogenously generated ROS is thought to be produced as a by-product of other reactions in the cell and is largely produced by the electron transport chain. In addition, nicotinamide adenine dinucleotide phosphate (NADPH) oxidase (NOX) isoenzymes produce H_2O_2 intentionally for signal transduction (Meitzler et al., 2013).

Mitochondria are known as the “power houses” of the cell, producing the majority (>80%) of total ATP formed from glucose through the process of oxidative phosphorylation (Figure 1.1). Oxidative phosphorylation is a multistep process that relies on the energy rich molecules nicotinamide adenine dinucleotide (NADH) and flavin adenine dinucleotide ($FADH_2$) that donate electrons in the reduction of molecular oxygen to water. This leads to several redox reactions that transport electrons across complexes I to IV. This electron transport is coupled to the transport of hydrogen ions across the mitochondrial membrane by the electron transport chain to generate a proton gradient and a transmembrane electrical potential that generates proton-motive force. The energy of the proton-motive force is used to power the ATP synthesis complex (also known as complex V) for the conversion of adenosine diphosphate (ADP) to yield ATP (Nickel et al., 2014). Under low levels of ADP, H^+ ions accumulate causing the electron transport chain to become more reduced, which allows electrons to escape and react with molecular oxygen leading to an increased formation of $O_2^{\bullet-}$ (Berg et al., 2002).

Complex I and Complex III are the major sites of ROS production (primarily superoxide and some hydrogen peroxide) in the electron transport chain. Complex III is the main producer of superoxide in both heart and lung mitochondria. However, complex I is the main source of superoxide production in the brain and is linked with the enhanced ROS production associated with several diseases, which are discussed in greater detail later section 1.2 (Kudin et al., 2005). The important roles that complex I and III play in the generation of ROS has been demonstrated with specific complex inhibitors. Administration of the complex I inhibitor rotenone and the complex III specific inhibitor antimomycin, have been shown in several models to lead to increased superoxide production and in the case of rotenone trigger apoptosis. Tumour necrosis factor alpha (TNF-

α) has also been shown to inhibit complex I and trigger apoptotic cell death pathways (Li et al., 2003).

Pharmacological compounds that induce members of the cytochrome P450 enzyme family have also been linked to the production of ROS. Phenobarbital administration has been shown to increase the expression of F2-Isoprostanes (F2-isoPs), an established biomarker of *in vivo* oxidative stress. Cytochromes P450 catalyse the oxygenation of substrates coupled with the simultaneous reduction of O_2, but this reaction has the potential to uncouple resulting in the production of ROS. However, phenobarbital has also been shown to decrease the expression of the antioxidant defence enzyme, glutathione peroxidase, which may account for the increase in ROS levels (Dostalek et al., 2008;Hrycay and Bandiera, 2015).

Other intracellular sources of ROS include: peroxisomes, lipoxygenases and cyclooxygenases. Peroxisomes carry out several key processes in eukaryotic cells including fatty acid β-oxidation to produce acetyl CoA, NADH and $FADH_2$, which results in the production of H_2O_2 as a bi-product (Poirier et al., 2006;Foerster et al., 1981). The oxidative metabolism of arachnoid acid mediated by lipoxygenases and cyclooxygenases also leads to the production of ROS (Cho et al., 2011).

Uncontrolled production of ROS leads to the modification of the structure and function of cellular proteins and lipids; and the modification of DNA. This can lead to cellular dysfunction: including impaired cell transport mechanisms, immune activation and inflammation, impaired energy metabolism, altered cell signalling and altered cell cycle; and ultimately result in the development of disease.

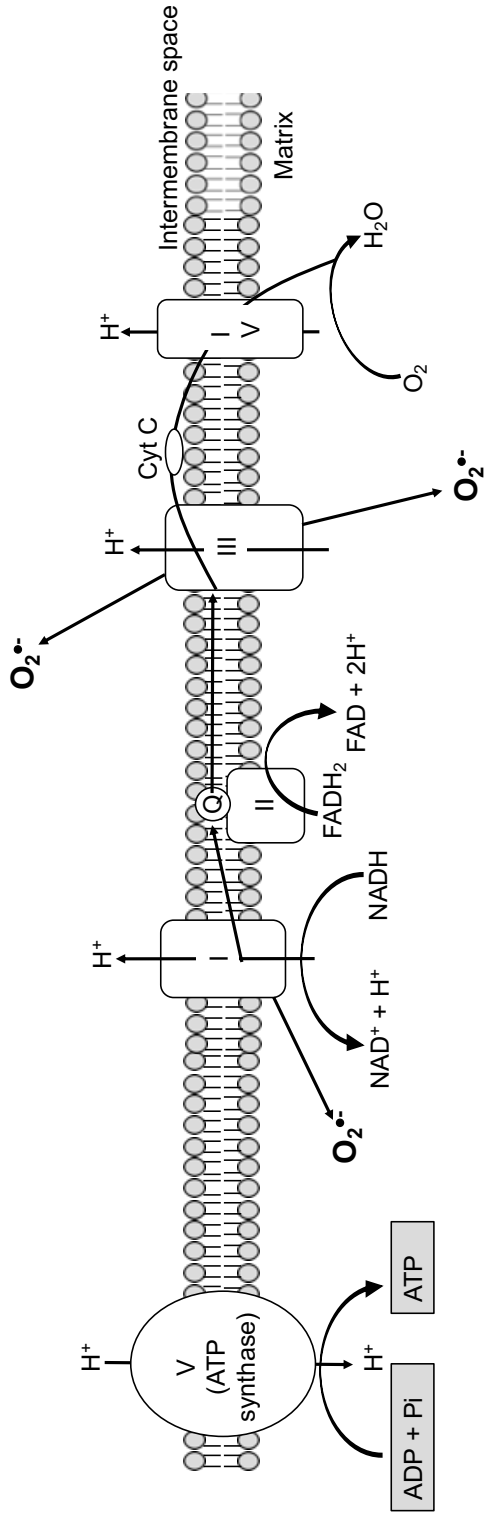


Figure 1.1: Overview of oxidative phosphorylation

NADH and $FADH_2$, provided from a variety of sources, and H^+ ions produced by the hydrolysis of these two molecules is used to generate the proton gradient required to produce ATP. Sources of ROS are indicated in bold font. Adapted from (Bratic and Larsson, 2013).

1.2 Role of oxidative stress in disease

It has been debated extensively whether the production of ROS is beneficial or deleterious for the cell and the impact that ROS has on cellular health maybe in part dictated by its levels. Cells are constantly producing reactive oxygen species; high levels of ROS are commonly linked with the development of disease. However, low levels of ROS have been shown to be pivotal to several physiological signalling pathways, such as MAPK signalling, through the modification of redox sensitive proteins which contain key cysteine residues that when modified change the protein into its active state. Tyrosine phosphatases are also cysteine rich proteins which under low levels of ROS (in particular hydrogen peroxide) are reversibly inactivated. This permits increased tyrosine phosphorylation of several proteins leading to their activation and the activation of several downstream signalling cascades that promote cell survival (Finkel, 2001; Finkel, 2011). Low levels of hydrogen peroxide have been shown to reversibly convert cystine residues that exist as thiolate anions to the sulfenic form, this process is pivotal to several signal transduction pathways and is thought to not be damaging to the cell. However, in the presence of high levels of hydrogen peroxide thionate anions are irreversibly converted sulfinic /sulfonic ions, which are extremely damaging to the cell (Schieber and Chandel, 2014).

The immune system provides a means of defence for cells against potentially damaging pathogens and impaired immune signalling, through either inhibition or hyperactivation of the immune system, is associated with the development of several diseases. Phagocytic neutrophils and macrophages exploit the destructive power of ROS during the wound healing process to kill invading pathogens. High levels of oxygen are needed during wound healing to compete with the increase ATP demand for tissue repair. Phagocytic neutrophils and macrophage cells engulf invading bacteria into vacuole like structures called phagosomes which have high levels of the enzyme NOX. Upon intake of oxygen, NOX isoenzymes in the phagosome reduce oxygen to either $O_2^{\bullet-}$ or H_2O_2 which in high levels kills the engulfed invading bacteria, in a process called “respiratory burst”. Whilst the respiratory burst is being carried out, high levels of H_2O_2 are produced and released into the surrounding environment, halting the growth of neighbouring bacteria and ensuring complete eradication of all pathogenic tissue preventing the formation of chronic wounds. Chronic Granulomatous disease

(GCD) occurs in patients harbouring mutations in NOX, this means they cannot reduce oxygen to produce ROS and therefore are subject to numerous infections and chronic wound formations (Halliwell, 2006;Dunnill et al., 2017).

The fact that the cell has a plethora of cell defences in place to combat ROS production indicates that these species may not be always beneficial to cell health but often pivotal to the development of disease. The balance between ROS production and elimination in the cell can be altered by either: deficiency in antioxidant defences, inhibition of complexes of the electron transport chain or exposure to xenobiotics. ROS production has been linked with a wide range pathological conditions such as neurological disorders (Parkinson's disease and Amyotrophic Lateral Sclerosis (ALS)), ageing, diabetes and cancer.

1.2.1 Neurological disorders and oxidative stress

The development of Parkinson's disease (PD) is thought to be linked to repetitive exposure to complex I inhibitors, which trigger $O_2^{\bullet-}$ production resulting in activation of the apoptotic cell death pathway. Mice that are administered rotenone have been shown to develop α-synuclein deposits in the brain, a classical hallmark of PD. These α-synuclein deposits have also been shown to induce ROS formation through inhibiting activating transcription factor 6 (ATF6), an important transcription factor involved in the unfolded protein response (UPR), leading to neurodegeneration (Turrens, 2003). Also, mutations in the mitochondrial targeting sequence (MTS), which is responsible for the translocation of the antioxidant manganese SOD (MnSOD) to the matrix, have been identified in PD suffers suggesting a genetic link between dysregulation of ROS production and the development of PD (Zeeshan et al., 2016).

ALS is an extremely debilitating muscular disorder of which the cause of 90-95% of cases is unknown. The remaining 5-10 % of all cases have a familial link, and 10% of these familial cases have been shown to have mutations in the gene encoding copper-zinc superoxide dismutase (CuZnSOD). These mutations lead to blockage of the transport of the enzyme into the inter membrane space and results in increase superoxide production (Turrens, 2003).

1.2.2 Ageing

Ageing has been linked to the build-up of DNA damage, which can be measured through the presence of 8-oxo-deoxyguanosine (8-oxo-dG), through

continuous exposure to ROS. Studies have shown that increases in age are associated with an increase in mitochondrial superoxide and hydrogen peroxide formation leading to increased oxidation of proteins. Deficiency in the mitochondrial enzyme required for DNA repair are found in patients suffering from Cockyanes syndrome, a disease characterised by premature ageing. Conversely, *drosphillia melanogaster* models have shown that increased expression of the ROS defence enzymes, SOD and catalase, is associated with increased life span. However, administration of antioxidants has not been conclusively shown to have any effects on aging in human models (Pickering et al., 2012; Liochev, 2013; Barja, 1999).

Human population studies have conclusively shown that women live longer than men. This increase in female life span maybe due to women having higher expression of the ROS defence enzyme, glutathione peroxidase. However, the exact mechanism behind this increase in glutathione peroxidase expression has not been determined but the female hormone oestrogen is thought to potential play a role (Viña et al., 2005).

1.2.3 Diabetes mellitus

High levels of oxidative stress have been shown to lead to the development of Diabetes mellitus (DM), a highly common metabolic and inflammatory disease. Diabetes can be divided into two subtypes that are called type 1 DM (T1DM) and type 2 DM (T2DM), which differ in their pathology but both exhibit impaired insulin signalling and high levels of ROS.

T1DM is often diagnosed in early life and is associated with a genetic risk factor. In T1DM immune cells attack healthy β-cells in the pancreas, which are responsible for the production of insulin, through ROS induced respiratory burst (Huang et al., 2016). T2DM is often associated with excessive nutritional intake leading to high blood glucose levels (hyperglycaemia) that is coupled with a lack of physical exercise. High glucose levels promote the formation of ROS through increasing the levels of diacylglyerol (DAG), which subsequently activates protein kinase C (PKC) resulting in phosphorylation of NOX isoenzymes leading to increased formation of ROS (Giacco and Brownlee, 2010). The increased ROS production associated with diabetes arises predominantly through the electron transport chain, but may also be due to increases in the activity of nitric oxide synthase and peroxidases and also toll like receptor 4 signalling. In diabetes, the

high levels of ROS lead to activation of the apoptotic death pathways which lead to inflammation and insulin resistance through altering insulin receptor downstream signalling. The rate of apoptosis in the cell is normally tightly coupled with rate of autophagy. However, impairments in autophagy have been shown in DM leading to an accumulation of unfolded proteins resulting in dysfunction of the endoplasmic reticulum (ER) and ER stress.

To further enforce the important role that oxidative stress plays in DM, alterations in the antioxidant defence system have been shown to be associated with this disease. For example, the levels of catalase and glutathione are decreased in T2DM patients (Newsholme et al., 2016;Volpe et al., 2018). Also, administration of antioxidants, such as mitoquionine, have been shown to inhibit β-cell death in cell models of DM (Zeeshan et al., 2016). Overall, these findings suggest that oxidative stress plays a fundamental role in DM development through increased ROS levels and impaired oxidative stress defences.

1.2.4 Cancer

Cancer is a multistep disease that has defined hallmarks associated with it. The hallmarks of cancer include: increased angiogenesis, sustained cellular proliferation and metastatic potential, and resistance to cell death pathways (Hanahan and Weinberg, 2011;Hanahan and Weinberg, 2000). ROS can play a role in the establishment of each of these hallmarks.

Cancer cells have enhanced oxygen and nutrient demand which requires additional blood flow. Angiogenesis is the formation of new blood vessels and is fundamental to cancer cell proliferation, and migration, and leads to life-threatening metastatic disease (Nishida et al., 2006). Low levels of oxygen, known as hypoxia, leads to the activation of the transcription factor hypoxia inducible factor-1 (HIF-1). HIF-1 is composed of two isoforms, designated HIF-1α and HIF-1β. The β isoform is constitutively active in all cells but the α isoform is regulated by ROS and activated upon exposure to H_2O_2 (Chandel et al., 2000;Chang et al., 2008). HIF-1α functions to inhibit the major regulator of angiogenesis, vascular endothelia growth factor (VEGF) (Ramakrishnan et al., 2014). VEGF has also been shown to activate NOX leading to ROS production, suggesting the presence of cycle of ROS stimulated angiogenesis in tumours. The requirement for ROS in the development of angiogenesis has been

demonstrated in both *in vivo* and *in vitro* models of the disease through the use of antioxidants and inhibitors of the mitochondrial electron transport chain. N-acetylcysteine (NAC), a potent antioxidant, and rotenone (a complex I inhibitor) have been shown to inhibit angiogenesis through the suppression of VEGF (Xia et al., 2007;Kim and Byzova, 2014;Liou and Storz, 2010;Galanis et al., 2008).

Metastasis occurs when cancer cells migrate through the blood circulation to proliferate in other locations in the body. The development of metastatic cancer dramatically decreases patient survival rate. For cells to migrate into the blood system they first must detach from the basal membrane. Metalloproteinases (MMPs), in particular MMP2 and MMP9, breakdown proteins of the extracellular matrix, freeing cells. High levels of ROS, caused by increased H_2O_2 production, activate MMPs and the expression of both MMP2 and MMP9 has been shown to be higher in several cancerous tissues in comparison to matched non-malignant control tissues (Roomi et al., 2009). Overexpression of the antioxidant defence enzyme MnSOD has been shown in rodent models to trigger cellular invasion (Connor et al., 2007). Also, high levels of mnSOD have been reported at the outer most edge of the tumour which is the location of the invading cells is and by contrast low levels of expression are found in the torpid centre of the tumour. This may be in part explained by the fact that mnSOD catalyses the conversion of superoxide to hydrogen peroxide which will stimulate the outer most edge of the tumour to invaded into the adjacent tissue (Zhang et al., 2002;Liou and Storz, 2010).

Errors in cell cycle are often implicated in the development of cancer. Cell cycle is an irreversible process controlled by cyclins and cyclin-dependant kinases (CDKs). Different combinations of cyclins and CDKs are activated and inhibited at different stages of the cell cycle. Once the cyclin has ensured the progression through to the next phase of the cycle they are degraded through ubiquitination, meaning that cell cycle is a unidirectional irreversible process. The activity of CDKs is controlled by the phosphatase, cdc25, and CDK-inhibitors (CKIs) (Barnum and O'Connell, 2014). Cdc25 stimulates cell cycle progression and is overexpressed in several cancer tissues. ROS decreases the activity of cdc25 through oxidising the highly reactive active site cysteine and through enhancing the degradation of the phosphatase (Rudolph J, 2005). ROS also

influences the ubiquitination of CKIs and can directly inhibit the proteasome (Verbon et al., 2012).

The role that ROS plays in apoptosis has been studied since the discovery that H_2O_2 exposure induces apoptosis (Pierce et al., 1991). Apoptosis, also known as programmed cell death, is controlled through the inhibition of anti-apoptotic proteins and the activation of pro- apoptotic proteins. There are two forms of apoptosis that take place in the cell, that are called extrinsic receptor-mediated apoptosis and intrinsic mitochondrial-associated apoptosis, and ROS plays a role in regulating both forms of apoptosis (Simon et al., 2000). Extrinsic receptor-mediated apoptosis involves a variety of signalling cascades. Nuclear factor–kappa beta (NF-κβ) is an anti-apoptotic protein that upon exposure to ROS is inhibited leading to apoptosis (Morgan and Liu, 2010). Also, apoptosis regulating kinase -1 (ASK-1) activity is controlled through ROS (Soga et al., 2012). ASK-1 is inhibited under low levels of ROS through an association with thioredoxin 1 (Trx1). Upon exposure to ROS several key thiols of Trx1 are oxidized leading to its dissociation from ASK-1. Activated ASK-1, activates c-Jun N-terminal kinases (JNK) and active JNK signalling leads to increased release of pro-apoptotic protein cytochrome C and the phosphorylation mediated inactivation of anti-apoptotic protein, B-cell lymphoma 2 (BCL-2) (Yamamoto et al., 1999). Increased mitochondrial ROS leads to triggering of the intrinsic apoptosis pathway leading to the formation of pores in the outer mitochondrial membrane, this triggers the translocation of cytochrome C from the intermembrane space to the cytoplasm. In the cytoplasm, cytochrome C binds to Apoptotic protease activating factor 1 (Apaf-1) and triggers the caspase cascade leading to cell death (El-Osta and Circu, 2016). Photodynamic therapy is often used for the treatment of breast and bladder cancer due to it being minimally invasive and with few associated side-effects. This therapy increases the levels of ROS, in particular highly cytotoxic singlet oxygen, through the interaction a specific wavelength of light, a photosensitiser and molecular oxygen (Jarvi et al., 2012).

1.3 Cellular defences against ROS

To combat the ever-present damage inflicted by ROS, the cell has developed an effective antioxidant defence system that comprises enzymatic (including; SOD, catalase (CAT) and glutathione peroxidase (GPX)) and non-enzymatic components (including; vitamin E, vitamin A, thioredoxin and glutathione) that serve to remove them.

1.3.1 Enzymatic antioxidants

Enzymatic antioxidants constitute the first line defence again ROS through suppressing or inhibiting ROS accumulation in cells (Ighodaro and Akinloye, 2017). SODs protect cells from the damaging and potent oxidizing effects of superoxide and limit the formation of hydroxyl radicals and nitric oxide. SODs function to covert two superoxide anions to molecular oxygen and H_2O_2, which can be later neutralised by catalases or glutathione peroxidases (Figure 1.2). SODs are a family of metalloenzymes that utilize metal ions as co-factors, such as; zinc (Zn), copper (Cu) and manganese (Mn), for the dismutation of superoxide. Family members not only vary in the identity of their associated metal co-factor but also their localization. Due to the large associated superoxide production, mitochondria express SOD isoenzymes. SOD2, also known as manganese superoxide dismutase (MnSOD), is expressed in the matrix of mitochondria. By contrast, CuZnSOD, also known as SOD1, is associated with the inner membrane of the mitochondria. Cytochrome C, located in the inner membrane space can also interact with superoxide leading to the generation of oxygen and the reduction of cytochrome C. Reduced cytochrome C alters the pH of the inter membrane space and leads to the spontaneous dismutation of superoxide. Extracellular SOD (ECSOD), also known as SOD3, is associated with the extracellular space (Azadmanesh and Borgstahl, 2018; Turrens, 2003; Ighodaro and Akinloye, 2017). Accumulation of hydrogen peroxide through the dismutation of superoxide carried out by SOD isoenzymes, can be damaging to the cell. Also, H_2O_2 can be converted to hydroxyl radicles through the Fenton reaction in the presence of iron (Figure 1.3). The Fenton reaction is commonly associated with cellular damage.

CAT is predominantly found in peroxisomes and functions to prevent the build-up of hydrogen peroxide by neutralizing it to molecular oxygen and water

by utilizing either iron or manganese as a co-factor (Shin et al., 2018). GPX isoenzymes are found in several cell compartments and function to neutralize H_2O_2 through the use of selenium as a co-factor coupled with the oxidation of glutathione. GPX is also involved in the conversion of lipid peroxidases to their corresponding alcohols and therefore limit lipid peroxidation. Since CAT is not expressed in mitochondria, GPX is responsible for the neutralization of all mitochondrial associated H_2O_2 (Ighodaro and Akinloye, 2017;Winterbourn, 2013). GPXs show a higher affinity for H_2O_2 than CAT, and under normal homeostatic conditions the majority of H_2O_2 is degraded by GPXs. However, when H_2O_2 is expressed at very high concentrations both GPXs and CAT are activated (Shin et al., 2018).

Mutations and deficiency in SOD, CAT and GPX isoenzymes have been implicated in a wide range of pathological disorders, highlighting the importance of enzymatic antioxidants in combating oxidative stress.

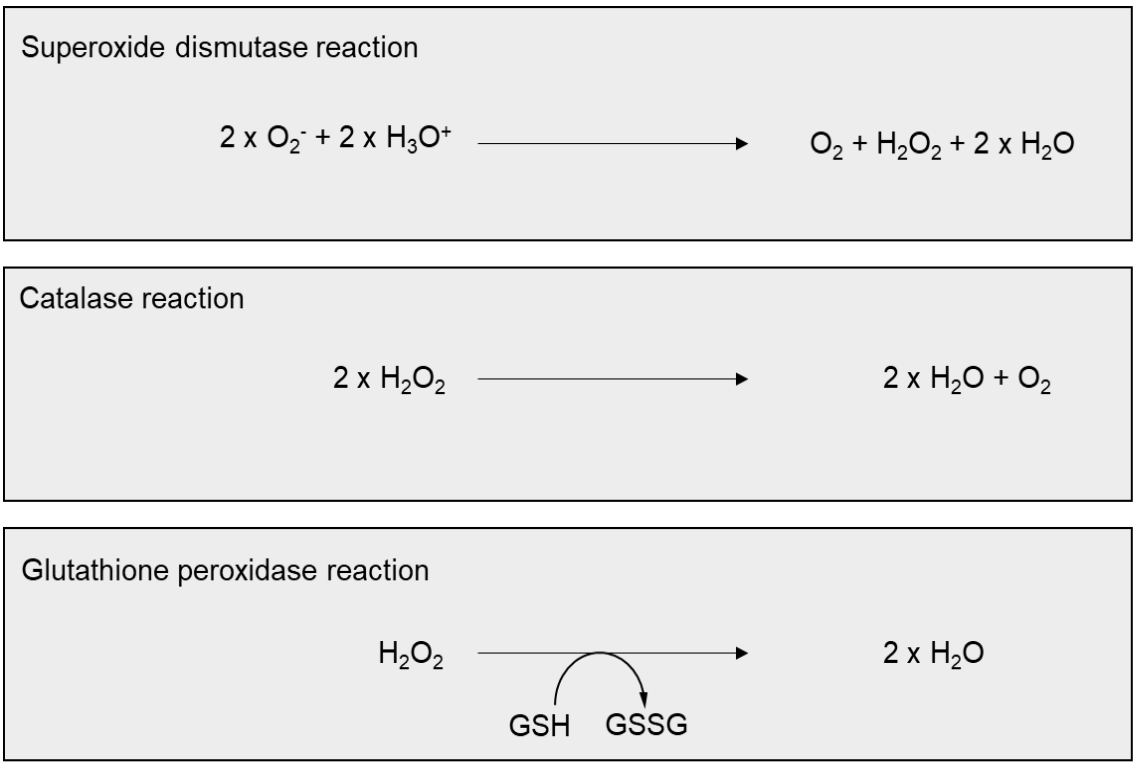


Figure 1.2: Reactions catalysed by enzymatic antioxidants

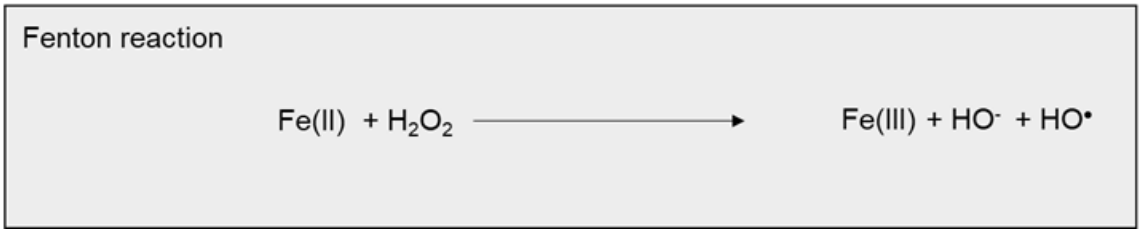


Figure 1.3: The Fenton reaction

1.3.2 Non-enzymatic antioxidants

Non-enzymatic antioxidants function as the second line defence against ROS through scavenging radicals via donating an electron (Ighodaro and Akinloye, 2017).

Vitamin C (also known as ascorbate) and vitamin E (also known as α-tocopherol) are important water-soluble antioxidants provided from dietary sources. Vitamin C also functions as a co-factor for several other enzymes. Lipid peroxidation is a potentially damaging free radical chain reaction that contributes to the development of several diseases and vitamin C protects against lipid peroxidation through scavenging ROS. By contrast, vitamin E is a potent peroxyl radical (ROO•) scanger that limits lipid peroxidation. The hydroxyl group of vitamin E interacts with the peroxyl radical leading to the formation of lipid hydrogen peroxide and the vitamin E radical (Vit E-O•), which can then be reduced back to vitamin E through interaction with vitamin C, a process called vitamin E recycling (Traber and Stevens, 2011; McCay, 1985; Traber and Atkinson, 2007). Due to their potent antioxidant abilities, both vitamin C and vitamin E are attractive therapeutics. However, clinical trials involving theses vitamins have had very disappointing and conflicting results (Gaziano et al., 2009).

In mammalian cells thiol dependant antioxidant systems exist that are based around; glutathione (see chapter 5 for more detailed description) and Trx, both of which are thought to function in parallel within the cell but also engage in cross-talk. Trx has several important functions in the cell that include: the direct scavenging of ROS, reducing key cysteine residues in transcription factors leading to enhanced DNA binding, supplying reducing equivalents for the DNA synthesis enzyme ribonucleotide reductase and functioning as an electron donner for thioredoxin peroxidases (Mustacich and Powis, 2000;Lu and Holmgren, 2014). The biological activity of Trx arises from key cysteine residues located in the catalytic site, these cysteine residues are oxidized upon interaction with the substrate and then converted back to the reduced form by the NADPH-dependant flavoenzyme, thioredoxin reductase (TrxR). There are distinct pools of Trx in the cell, the cytosolic pool utilizes Trx1 and TrxR1 isoforms and the mitochondrial pool uses Trx2 and TrxR2 isoforms (Collet and Messens, 2010).

1.4 The NRF2 antioxidant defence system

The ubiquitously expressed transcription factor, Nuclear factor erythroid 2-related factor 2 (NRF2) is regarded as the master regulator of cellular oxidative stress defences through its ability to regulate the basal and inducible expression of wide range of target genes, some of which encode proteins with antioxidant activities. NRF2 is thought to be involved in the trans-activation of over 200 target genes, the majority of which are cytoprotective (Rojo de la Vega et al., 2018).

1.4.1 Discovery of NRF2

NRF2 was first identified in 1994 through a screen of a λgt 11 cDNA expression library isolated from K562 cells using a tandem repeat recognition probe with an extended activator protein -1 (AP-1) sequence (Moi et al., 1994). One of the clones that was isolated was identified as a ubiquitously expressed 66 kDa protein with a cap'n'collar basic-region leucine zipper (CNC-bZIP) binding domain in its C-terminus and an acidic activator domain in the N-terminus. These researchers, gave this protein the name NRF2, the third member of the CNC-bZIP family of proteins to be discovered.

1.4.2 CNC-bZIP family members

Members of the CNC-bZIP family include the following: *caenorhabditis elegans* (*c.elegans*) skinhead-1 (SKN-1) protein, *drosophila melanogaster* cap'n'collar protein (CncC), nuclear factor-erythroid 2 p45 subunit (NF-E2 p45), nuclear factor erythroid 2-related factor 1 (NRF1, gene name *NFE2L1*), NRF2 (gene name *NFE2L2*), nuclear factor erythroid 2-related factor 3 (NRF3, gene name *NFE2L3*) and related proteins BTB and CNC homology 1(BACH1) and BACH2. Family members are characterized according to the presence of a 43-amino acid homology region called the cap'n'collar (CNC) domain, located immediately at the N-terminus of the basic DNA binding domain.

NF-E2 p45, expressed only in haematopoietic tissues, is the founding member of the CNC-bZIP family of proteins. The bZIP domain of NF-E2 p45 shows extensive homology with both the drosophila CNC protein and the *c.elgans* protein SKN-1, and is thought to be involved in the activation of transcription (Kotkow and Orkin, 1995). In mice, knock out (KO) of NF-E2 p45 causes extensive haemorrhaging, as it regulates several critical genes involved in platelet production (Shivdasani et al., 1995).

Family members NRF1 and NRF2 are both ubiquitously expressed, control the expression of antioxidant response element (ARE) containing genes and both dimerize with small musculoaponeurotic fibrosarcoma (Maf) proteins (Bugno et al., 2015). However, there are several differences between the two CNC-bZIP family members. NRF1 KO in mice is embryonically lethal (Chan et al., 1998), whereas NRF2 KO mice are viable with no adverse phenotype (Chan et al., 1996). This highlights that these two transcription factors cannot compensate for each other. NRF1 is either associated with the ER membrane though its N-terminal domain (NTD), a domain which is absent from NRF2 (Figure 1.4), or in the nucleus in its active form. For NRF1 to be activated it requires deglycosylation of several Asn residues, followed by proteolytic cleavage, this process is not required for NRF2 to be activated (Zhang et al., 2007b). Both NRF1 and NRF2 are regulated at a transcriptional and a post-translational level. The levels of NRF2 are principally controlled through association with the cytoplasmic negative regulatory protein, Kelch-like ECH-associated protein 1 (KEAP1), resulting in 26S proteasomal degradation of the transcription factor (this will be discussed in detail later). NRF1 can also be bound by KEAP1, however this does not lead to the degradation of NRF1 (Zhang et al., 2006b). NRF1 has been shown to degraded through interactions with the F-Box and WD repeat domain-containing 7 (FBW7) protein (Biswas et al., 2011), the beta-transducin repeat containing (β-TrCP) complex and the endoplasmic reticulum associated degradation (ERAD) ligase, Hrd1 (Tsuchiya et al., 2011). The transcriptional activity of NRF1 has also been shown to be influenced by both glycogen synthase kinase-3 (GSK-3) (Biswas and Chan, 2012) and casein kinase 2 (CK2) (Tsuchiya et al., 2013).

NRF3, unlike NRF1 and NRF2, is not ubiquitously expressed but is highly expressed in placenta and B-cells (Kobayashi et al., 1999). NRF3 KO mice are viable with no adverse phenotype (Derjuga et al., 2004). NRF3, like NRF1, due to its NTD is sequestered to the ER and must undergo deglycosylation of Asn residues and proteolytic cleavage for activation (Nouhi et al., 2007; Zhang et al., 2009). Like both NRF1 and NRF2, NRF3 dimerises with small Maf proteins and binds to target genes containing ARE sequences. Also, NRF3 has been shown to be degraded by Hrd1 and β-TrCP (Chowdhury et al., 2017; Chevillard and Blank, 2011).

There is thought to be significant overlap in the ARE-containing genes that are trans-activated by NRF1 and NRF2. For example several genes involved in fatty acid metabolism have been shown to regulated by both NRF1 and NRF2 (Tsujita et al., 2014). Due to their diverse subcellular localizations it is thought that NRF1, NRF2 and NRF3 may work in concert to mediate oxidative stress defences in the cell.

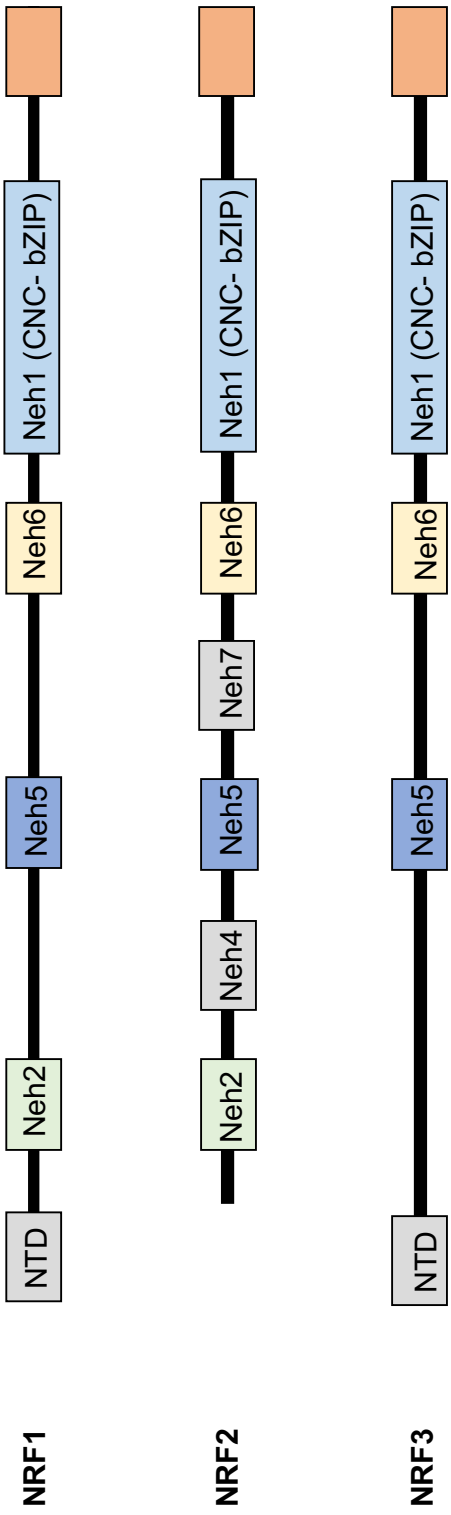


Figure 1.4: Domain structure of the three related CNC-bZIP transcription factors NRF1, NRF2 and NRF3.

1.4.3 Structure of NRF2

Mouse NRF2 is composed of 597 and human NRF2 of 605 amino acids. NRF2 can be divided into seven highly conserved protein domains, called NRF2-ECH homology (Neh) domains 1 – 7 (Figure 1.5). This nomenclature was chosen due to the extensive shared sequence identity of human NRF2 with the chick orthologue protein, erythroid cell-derived protein with CNC homology (ECH). Each of these seven Neh domains has a distinct function and confers a unique element to overall NRF2 signalling (Tebay et al., 2015; Itoh et al., 1995).

The Neh1 domain of NRF2 contains the CNC-bZIP region, which was first characterised in *drosophila melanogaster* as a DNA-binding region, and is critical for the activity of the transcription factor and its ability to dimerize (Mohler et al., 1991). The Neh2 domain has been demonstrated to be the site interaction with KEAP1 and is necessary for KEAP1-dependent degradation of NRF2 protein (Itoh et al., 1999). The most C-terminal domain of NRF2, Neh3, has been shown to function as a trans-activation domain and interacts with chromodomain helicase DNA binding protein 6 (CHD6) (Nioi et al., 2005). Both the Neh4 and Neh5 domains have been shown to function as transactivation domains that are required for the transcriptional activity of NRF2 (Zhang et al., 2007a). These domains also function to cooperatively bind the co-activator CREB binding protein (CBP) (Katoh et al., 2002). The Neh5 has also been demonstrated to be the interaction site with chromatin remodelling protein, BRG1 (gene name *SMARCA4*) (Zhang et al., 2006a) and the mediator subunit 16 (MED16) (which also binds to the Neh1 domain) (Sekine et al., 2016). The Neh6 domain is the site of KEAP1-indepedent regulation of the transcription factor through β-TrCP mediated degradation and the site of GSK-3 phosphorylation (McMahon et al., 2004;Chowdhry et al., 2013;Rada et al., 2011). This form of NRF2 regulation will be discussed in detail later in Chapter 3. The Neh7 domain is the site of interaction with retinoid x receptor alpha (RxRα), which leads to inhibition of ARE-driven target gene expression (Wang et al., 2013).

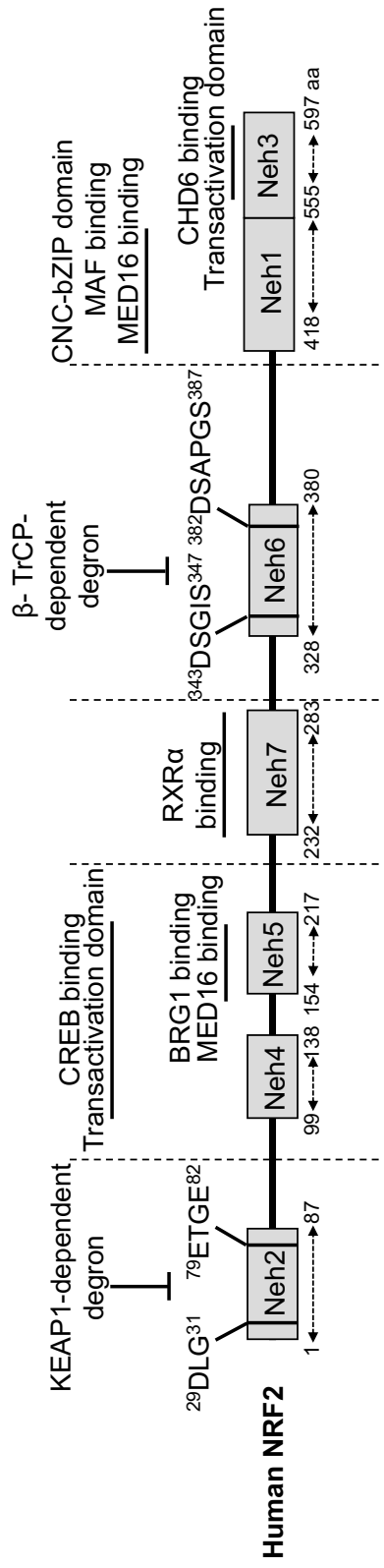


Figure 1.5: Detailed structure of human NRF2

Each of the seven Neh domains within NRF2 are indicated on the figure in the grey boxes. The size of each domain is indicated below the figure between the stripped arrows. The function/ interaction of each domain is highlighted above the schematic.

1.4.4 Antioxidant response element (ARE)

A common feature in all NRF2-target genes is the presence of a cis-acting DNA sequence that is responsive to oxidative stress in the promoter region, termed the ARE sequence. The ARE sequence was initially discovered in the early 1990's through various mutational and deletion studies of the genes encoding the drug detoxification enzymes glutathione S-transferase alpha-2 (*GSTA2)* and NAD(P)H dehydrogenase quinone 1(*NQO1*) (Rushmore and Pickett, 1990; Wang and Williamson, 1996; Favreau and Pickett, 1993; Rushmore et al., 1991). The significance of ARE sequences has been further highlighted through the use of various chemo-preventive agents such as butylated hydroxyanisole (BHA) and ethoxyquin, which are thought to impart their anti-cancer effects through inducing the expression of drug detoxification genes (Hayes et al., 2000). ARE sequences have now been identified in over 200 genes, however it should be noted that ARE sequences show substantial variation (Nioi et al., 2003), genes can have multiple ARE sequences and not all ARE sequences are functional (Wasserman and Fahl, 1997). The core ARE sequence is often described as 5'-TGACNNNGC-3', where N represents any amino acid (Nioi et al., 2003; Tebay et al., 2015).

1.4.5 NRF2 mechanism of action and interaction with small MAF proteins

NRF2 activity is dampened under normal homeostatic conditions due to the association with KEAP1 (this is discussed in detail in 1.5.2.2). Under conditions of oxidative stress or upon exposure to thiol-active electrophilic compounds, KEAP1 is altered in such a way that impairs its ability to ubiquitinate NRF2. This permits the stabilisation and translocation of NRF2 into the cell nucleus. Once in the cell nucleus NRF2 binds to ARE-sequences present in the promoter of NRF2-target genes (discussed in 1.4.3) in combination with small Maf (sMaf) proteins, resulting in active gene transcription.

The sMaf family of transcription factors is composed of three members, called; MafG, MafF and MafK, which show no functional differences or differential expression patterns but can heterodimerize or homodimerize with themselves or members of the CNC bZIP family (Motohashi et al., 2004; Yoshida et al., 2005). The leucine zipper domain of sMaf proteins have been shown to hetrodimerize with the bZIP region in the Neh1 domain of NRF2. This interaction has been shown to influence the expression of several ARE-containing genes (Toki et al.,

1997; Itoh et al., 1997). It is potentially not surprising that sMaf protein have been shown to interact with ARE sequence found in NRF2-target genes as this sequence closely resembles the Maf recognition element sequence, TGCTGAC(G)TCAGCA (Katsuoka et al., 2005; Yoshida et al., 2005). Small Maf protein expression is tightly regulated with high expression levels being associated with transcriptional repression. Oxidative stress (predominantly exposure to H_2O_2) and heat shock have both been shown to alter the expression pattern of small Maf proteins and therefore can regulate the expression of NRF2-target genes (Blank, 2008).

1.4.6 Genes regulated by NRF2

To date there are over 200 validated NRF2-target genes that can be described broadly as cytoprotective. The control of such a large cohort of genes demonstrates the influence that NRF2 signalling plays in the cell. The development of NRF2 KO mice and potent NRF2 inducers has been key to the identification of several NRF2-target genes (Chorley et al., 2012; Thimmulappa et al., 2002; Hu et al., 2006). NRF2 was first implicated in regulation of the expression of target genes involved in drug detoxification, such those encoding glutathione s-transferase (GST) isoenzymes and NQO1. Now it its thought that NRF2 controls the expression of an array of target genes that can be broadly divided into the following twelve subgroups: drug detoxification, glutathione-based antioxidants, thioredoxin-based antioxidants, pentose phosphate pathway and NADPH regeneration, purine synthesis, iron sequestration, stress markers, proteasome subunits, serine/glycine biosynthesis, apoptotic proteins, autophagy-related proteins and oxidative stress defences (a list of some NRF2-target genes that fall into theses twelve subgroups is given in Table 1.1.). Some differences exist between the magnitude of induction of certain NRF2-target genes between mice and humans. For example, in humans NRF2 robustly regulates the expression of certain aldo-keto reductase (AKR) family members but not GST isoenzymes, however the converse is seen in mice. By contrast *NQO1* expression is highly induced upon activation of NRF2 in both humans and mice (Tebay et al., 2015).

Due to the extensive homology between NRF1 and NRF2, it was originally thought that both transcription factors could potentially control the same target genes. However, NRF1 KO mice are embryonically lethal, whereas NRF2 KO

mice are viable, suggesting that the two transcription factors have different functions. Generation of a hepatocyte specific NRF1 KO mice revealed that a loss of NRF1 does not affect the expression of prototypic NRF2 target genes but decreases the expression of ARE- containing target genes involved in stress reposes, glycolysis and transport of xenobiotics (Ohtsuji et al., 2008).

Biochemical function	Short gene name	Full description
Drug detoxification	*ABCB6*	ATP-binding cassette, subfamily B (MDR/Tap) member 6
	AKR1B10	Aldo-keto reductase family 1, B10
	AKR1C1	Aldo-keto reductase family 1, C1
	AKR1C3	Aldo-keto reductase family 1, C3
	NQO1	NAD(P)H: quinone oxidoreductase 1
NFE2L2/ *KEAP1* expression	*NFE2L2*	Nuclear factor (erythroid-derived 2)-like 2
	KEAP1	Kelch-like ECH-associated protein 1
Antioxidants: GSH-based	*GCLC*	Glutamate-cysteine ligase catalytic subunit
	GCLM	Glutamate-cysteine ligase modifier subunit
	GPX2	Glutathione peroxidase 2
	SLC7A11	Sodium-independent cysteine-glutamate antiporter
Antioxidants: TXN-based	*PRDX1*	Peroxiredoxin 1
	PRDX6	Peroxiredoxin 6
	SRXN1	Sulfiredoxin 1
	TXNRD1	Thioredoxin reductase 1
	TRDX	Thioredoxin
Pentose phosphate pathway (PPP) enzymes and NADPH generation	*G6PD*	Glucose-6-phosphate dehydrogenase
	IDH1	NADP-dependent isocitrate dehydrogenase
	ME1	Malic enzyme 1
	PGD	6-phosphogluconate dehydrogenase
	TALDO1	Transaldolase
Purine synthesis	*MTHFD2*	Methylenetetrahydrofolate dehydrogenase
	PPAT	Phosphoribosyl pyrophosphate amidotransferase
Iron signalling	*FTH1*	Ferritin heavy polypeptide
	FTL	Ferritin light polypeptide
	HMOX1	Heme oxygenase 1
Proteasome subunits	*PSMA1*	Proteasome subunit alpha type-1
	PSMA4	Proteasome subunit alpha type-4
	PSMB5	Proteasome subunit beta type-5
	PSMC1	Proteasome AAA-ATPase subunit Rpt2
Serine/glycine biosynthesis	*PHGDH*	Phosphoglycerate dehydrogenase
	PSAT1	Phosphoserine aminotransferase 1
	SHMT2	Serine hydroxymethyltransferase 2
Apoptotic proteins	*BCL2*	B-cell lymphoma 2
Autophagy	*SQSTM1*	Sequestosome 1 (p62)
	ATG7	Autophagy protein 5
	ATG5	Autophagy protein 7
	CALCOCO2	Calcium binding and coiled – coil domain 2
	ULK1	UNC-51 autophagy activating kinase 1
	LC3B	Microtubule-associated protein 1A/1B-light chain 3 B (LC3B)
	GABARAPL1	Gamma-aminobutyric acid receptor-associated protein

Table 1.1: A list of genes regulated by NRF2 divided in different subgroups according to their function.

1.5 Molecular regulation of NRF2

As NRF2 controls a plethora of ARE-containing target genes involved in a wide range of cellular processes and so several forms of molecular regulation exist to control the levels and the activity of the bZIP-CNC transcription factor in the cell. NRF2 is regulated at both a transcriptional and post-translational level and dysregulation of these molecular regulation events is often implicated in the development of disease.

1.5.1 Transcriptional regulation of the *NFE2L2* gene

The ability to upregulate expression of drug-metabolizing enzymes is critical for cell survival upon exposure to stress. Several genes encoding drug-metabolizing enzymes contain either ARE or xenobiotic response element (XRE) DNA sequences in their promoters. The transcription factor aryl-hydrocarbon receptor (AhR) binds to XRE sequences whereas NRF2 binds to ARE sequences. It was originally thought that these pathways were completely independent of each other but it has now been revealed that there is some degree of cross-talk between the two (Hayes et al., 2009). Multiple XRE sequences have been identified in NRF2 (Miao et al., 2005) which can be induced by AhR and the effect of compounds that induce ARE-driven gene expression has been shown to be reliant on the presence of AhR (Miao et al., 2004). Contrastingly, the induction of several AhR-target genes has been shown to be dependent on NRF2 (Yeager et al., 2009; Ma et al., 2004). Also, though deletion and mutagenesis analysis it has been revealed that *NFE2L2* contains two weak ARE-like sequences in its promoter region thereby enabling NRF2 to autoregulate its own expression (Kwak et al., 2002).

Additionally, *NFE2L2* contains a TPA-response element (TRE) within the ARE site which allows increased *NFE2L2* transcription by the Kirsten rat sarcoma (*KRAS*) gene (Tao et al., 2014; Raghunath et al., 2018; DeNicola et al., 2011).

1.5.2 Post-translational regulation of NRF2

NRF2 is principally controlled at a protein stability level, with low and high levels of NRF2 being detrimental to cellular health. In the cell, NRF2 is constantly being synthesized and degraded to allow a pool of rapidly available NRF2 to mediate the rapid cellular response upon detection of stress and attenuation of its turn-over. There are two main routes by which NRF2 is post-translationally

regulated, namely, KEAP1-dependent regulation and KEAP1-independent regulation, both of which result in 26S proteasomal degradation of the transcription factor.

1.5.2.1 KEAP1-dependent degradation of NRF2

Under homeostatic conditions ARE-containing gene expression is low due to KEAP1 impeding NRF2 transcriptional activity through controlling both the subcellular localization and rapid degradation of the transcription factor (Itoh et al., 2003).

KEAP1 is a member of the BTB-Kelch family of proteins that functions as a substrate adaptor protein for the cullin 3 (CUL3) RING-box 1 (Rbx1) E3 ubiquitination ligase complex (designated CRL^{KEAP1}) to mediate 26S proteasomal degradation of target proteins (Pintard et al., 2004; Zhang et al., 2004). KEAP1 is homodimer composed of five domains. As shown in Figure 1.6, the domains in KEAP1 are: the N-terminal region (NTR), a broad complex, tram-track, bric-a-brac (BTB) domain which is responsible for the critical dimerization of KEAP1 (Zipper and Mulcahy, 2002) and binding to CUL3, the cysteine rich intervening region (IVR), a Kelch-repeat domain which contains six Kelch-repeats that form a six-bladed β-propeller structure (Li et al., 2005; Canning et al., 2015). The Kelch repeat domain of KEAP1 binds to the amino-terminal Neh2 domain of NRF2 which was originally thought to lead to impaired nuclear translocation of the transcription factor and thus decreased NRF2-target gene expression (Itoh et al., 1999). However, it is now widely excepted that the binding of KEAP1 to the Neh2 domain leads to the ubiquitination and proteasomal degradation of NRF2 (Itoh et al., 2003; McMahon et al., 2003). Several models have been proposed to explain the interaction between KEAP1 and NRF2, with the most widely excepted model being the two-site binding mechanism of the ‘hinge and latch’ model (also known as the ‘two-site tethering’ mechanism) (Tong Kit et al., 2006; McMahon et al., 2006). This model is founded on that concept that two subunits of KEAP1 bind to one subunit of NRF2 through each KEAP1 subunit within the dimer binding to a different motif in the Neh2 domain (Figure 1.7). Importantly, it has also been proposed that the KEAP1 dimer binds first to a high-affinity ETGE motif (hinge) and then subsequently to a 100-fold lower-affinity DLG motif (latch) (Tong et al., 2006). Mutagenesis studies have revealed that the ETGE motif is essential for KEAP1 interaction with the Neh2 domain and the turnover of NRF2 (Kobayashi

et al., 2002;McMahon et al., 2003). Moreover, the bottom face of the β-propeller structure of KEAP1 is positively charged due to the presence of highly conserved Arg residues, which may form an electrostatic interaction with the acidic ETGE motif (Li et al., 2005). The Neh2 domain also contains a cluster of seven highly conserved Lys residues that have been shown to be critical for KEAP1-mediated ubiquitination of NRF2 (Tong et al., 2006;McMahon et al., 2006). Interaction of KEAP1 with cytoskeletal actin filaments has been shown to provided scaffolding that is essential for KEAP1 function (Kang et al., 2004).

Under conditions of oxidative stress or upon exposure to electrophilic/thiol based compounds, several key cysteine residues in KEAP1 become modified and NRF2 can no longer be ubiquitinated (Itoh et al., 2003). NRF2 becomes locked in a conformation bound to KEAP1 in the cytoplasm. This leads to the accumulation and nuclear translocation of newly synthesized NRF2 as it can bypass the CRL^{KEAP1} complex. The fate of KEAP1 after cysteine modification is still under debate with some groups suggesting that KEAP1 becomes the target of CUL3 ubiquitination followed by degradation through a proteasome-independent route (Zhang et al., 2005) and others suggesting it is degraded though autophagy (Taguchi et al., 2012).

KEAP1 is regarded as a cysteine-rich protein with an average cysteine content of 4%, which is double that found in most proteins (Hansen et al., 2009; Tebay et al., 2015). There are 27 cysteine residues present in human KEAP1 and 25 in the mouse protein. Interestingly, several of these cysteines are located next to positively charged amino acids which maintains them in a thiolate anion state. These cysteines residues have been extensively demonstrated to be the sensors of a variety of inducers leading to the hypothesis of the cysteine code, which hypothesises that different inducing compounds will modify different combinations of cysteine residues to modulate the NRF2 stress response pathway (Dinkova-Kostova et al., 2017). The best characterized and most studied cysteine is Cys-151 which is located in the BTB domain of KEAP1 (Figure 1.8) and has been shown to be the sensor for sulforaphane (SFN) (Zhang et al., 2004;Zhang and Hannink, 2003), nitric oxide (McMahon et al., 2010), CDDO-imidazolide (CDDO-Im) (Saito et al., 2016) and tert-butylhydroquinone (t-BHQ) (Takaya et al., 2012). Modification of Cys-151 by all three of these electrophiles or mutations of this site to a Ser residue leads to the activation of NRF2 through

inhibition of KEAP1 (Zhang et al., 2004). Cys-273 and Cys-288 are in the IVR domain and are essential for the ubiquitination of NRF2 (Zhang and Hannink, 2003). Cys-273 can sense prostaglandins that are Michael reaction acceptors and Cys-288 can sense alkenals (McMahon et al., 2010; Kobayashi et al., 2009; Saito et al., 2016). Also, endogenous reactive metals, such as zinc (Zn^{2+}), increase in the presence of chemical stressors and lead to activation of the NFR2-signaling pathway through a non-covalent interaction with His-225,Cys-226 and Cys-613 (McMahon et al., 2010). ROS can also directly interact with cysteine residues of KEAP1 to stimulate NRF2 activity. Hydrogen sulphide (H_2S) and H_2O_2 have been shown to directly modify Cys-613 and Cys-226, inhibiting KEAP1 activity (Hourihan et al., 2012; Fourquet et al., 2010). Also, several enzymes, such as PTEN, PARKIN, GAPDH contain highly reactive cysteine residues in there activate site, meaning that their activity can be strongly influenced by ROS levels (Go et al., 2015).

Work by Ben Major and colleagues has revealed that KEAP1 not only binds NRF2 but can function as a substrate adaptor protein for several other proteins. These other KEAP1-binding proteins can function as indirect regulators of NRF2 activity through outcompeting NRF2 for KEAP1 binding, resulting in enhanced nuclear NRF2; this concept is discussed in more detail in Chapter 5 (Hast et al., 2013).

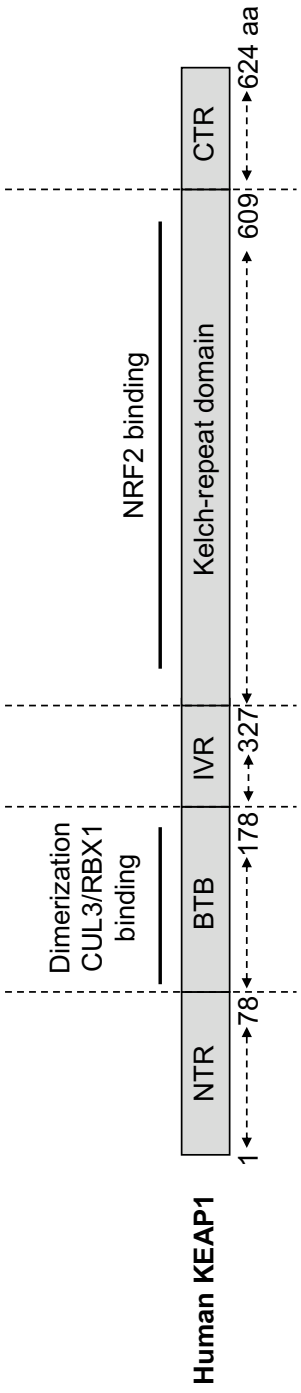


Figure 1.6: Structure of human KEAP1 protein

Different domains of KEAP1 are indicated in the light grey boxes. The size of each domain is highlighted below the figure between the striped arrows. The function/interaction of each domain is highlighted above the schematic.

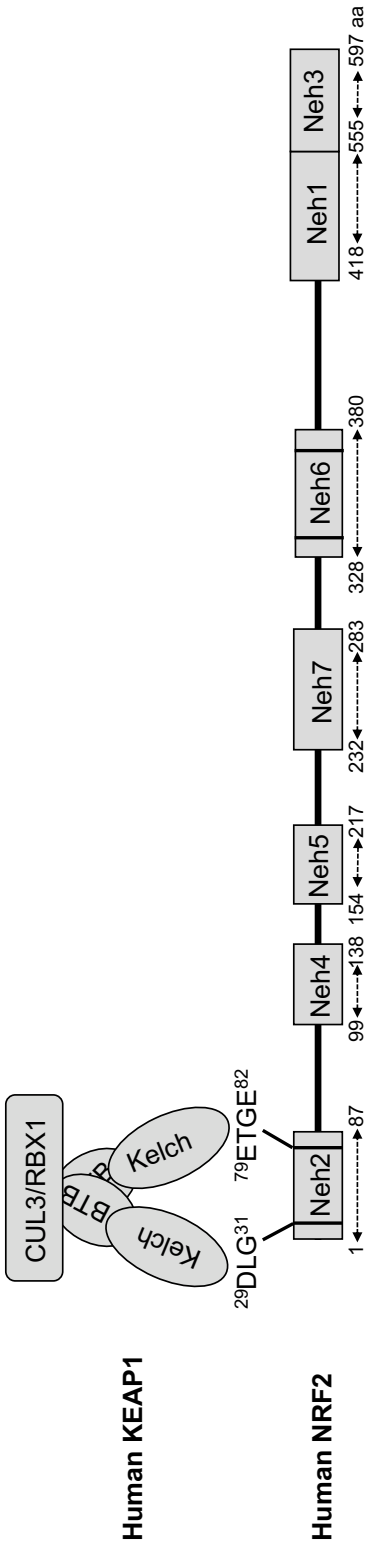


Figure 1.7: Hinge and latch model of the interaction between KEAP1 and NRF2 binding.

The DLG and ETGE motif sequences in the Neh2 domain to which KEAP1 binds are indicated above the schematic of human NRF2. KEAP1 is represented as a dimer; of which the Kelch domain binds NRF2 and the BTB domain interacts with CUL3/RBX1.

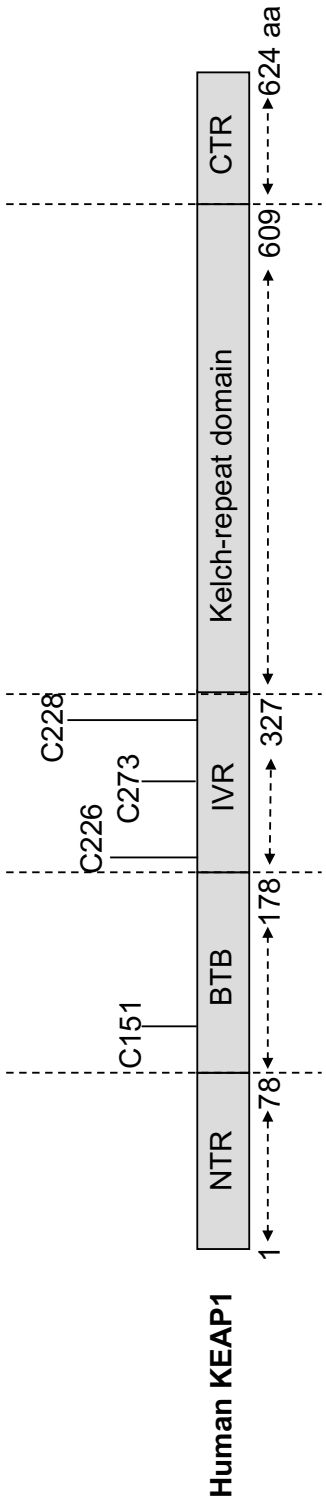


Figure 1.8: Location of four reactive Cys residues found within human KEAP *The domains of human KEAP1 are shown in the grey boxes. The size of each domain is indicated below the figure between the striped arrows. The locations of four reactive Cys residues (C) are indicated above the schematic.*

1.5.2.3 KEAP1–independent degradation of NRF2

As discussed above, KEAP1 is responsible for the cytoplasmic regulation of NRF2 through stimulating ubiquitin-mediated degradation of the transcription factor. The nuclear regulation of NRF2 is considerably less well characterised and is mediated by a separate KEAP1-independent degradation route involving GSK-3 and the $SCF^{\beta\text{-}TrCP}$ ubiquitin complex; this will be discussed in detail in Chapter 4.

1.5.3 Metabolic regulation of NRF2 by AMPK and mTOR

AMP-activated protein kinase (AMPK) functions to regulate several metabolic processes in the cell, such as proliferation, lipid metabolism, glucose metabolism and autophagy, which are also influenced by NRF2-signaling suggesting that there is a degree of cross-talk between theses pathways. AMPK is a serine/ threonine protein kinase that functions as an energy sensor and is activated under low levels of ATP to stimulate the production and decrease the consumption of ATP in the cell (Herzig and Shaw, 2018). Several groups have shown that AMPK directly interacts with NRF2 through phosphorylation of the CNC-bZIP protein at either Ser-40 or Ser-550, leading to enhanced nuclear accumulation of the transcription factor and increased expression of the NRF2-target genes, *NQO1* and *HMOX1* (Mo et al., 2014;Joo et al., 2016).

NRF2 activity is also influenced by the mechanistic target of rapamycin (mTOR) pathway which dictates the activity of both NRF2-regulatory proteins KEAP1 and β-TrCP. The mTOR protein is a serine/threonine kinase that is exists in two structurally and functionally distinct protein complexes called; mTORC1 and mTORC2, which function to regulate protein synthesis and autophagy. Hence alterations in the mTOR pathway are implicated in the development of several diseases, such as cancer which often have associated aberrant mTOR signalling (Laplante and Sabatini, 2012). The mTORC1 pathway regulates KEAP1 activity through phosphorylation of the autophagy adaptor protein, p62 at Ser-351 (Ichimura et al., 2013). Phosphorylation of p62 increases its affinity towards KEAP1 leading to KEAP1 degradation through autophagy and enhanced nuclear NRF2. The mTORC2 pathway influences NRF2 degradation through inactivating GSK-3 via phosphorylation of AKT at Ser-473, a site critical for AKT activity (Sarbassov et al., 2005). Inactivation of GSK-3 impairs the formation of

the phosphodegron motif to which the substrate adaptor protein β-TrCP binds, leading to enhanced NRF2 activity.

1.5.4 Other proteins that regulate NRF2

Regulation of NRF2 by KEAP1 and β-TrCP are by far the most well studied forms of regulation. However, there are additional mechanisms that alter NRF2 activity but their overall contribution to NRF2 signalling is currently not well characterized.

For example, the E3 ubiquitin ligase Hrd1, which is involved in the ERAD pathway, has been shown in the context of liver cirrhosis to ubiquitinate NRF2 leading to the degradation of the CNC-bZIP protein (Wu and Rapoport, 2018; Wu et al., 2014). Acetylation is an important cellular process that involves the introduction of functional acetyl groups which can alter protein function. Multiple Lys residues in the DNA-binding Neh1 region have been shown to undergo acetylation by p300/CBP, under arsenite-induced oxidative stress (Sun et al., 2009). These acetylation events are required for maximal binding of NRF2 to ARE sequences present in the promoters of NRF2-target genes and hence acetylation increases NRF2-target gene expression. Conversely the deacetylase sirutin 1 (SIRT1) has been shown to decrease NRF2 acetylation, leading to impaired ARE binding and reduced NRF2-target gene expression (Kawai et al., 2010).

Under conditions of hypoxia, NRF2 is regulated through a KEAP1-independent route involving seven in absentia homolog 2 (SIAH2). SIAH2 is a key regulator of the hypoxic response that binds directly to NRF2 stimulating the proteasomal degradation of the transcription factor (Baba et al., 2013; Baba and Miyazaki, 2016).

Also, NRF2 can undergo polysumoylation by small ubiquitin-related modifier 1 (SUMO1) and SUMO2, which leads to the translocation of NRF2 to the nucleus. Polysumoylated NRF2 is then targeted by the E3 ubiquitin ligase, RING-finger protein 4 (RNF4) leading to its degradation (Malloy et al., 2013).

1.6 The role of NRF2 in disease

The levels of NRF2 are tightly regulated in the cell, with both low and high levels of the transcription factor contributing to the development of disease; high levels of NRF2 are associated with the promotion of cancer; and low levels of NRF2 are associated with Parkinson's disease. Multiple mechanisms may lead

to abnormal NRF2 expression such as: mutations in *KEAP1* and *NFE2L2* (the gene that encodes NRF2 itself) (Huppke et al., 2017; Taguchi and Yamamoto, 2017), hypermethylation of the *KEAP1* promoter (Hanada et al., 2012) and alterations in KEAP1 activity through modification of cysteine residues. The role that NRF2 plays in the development of Parkinson's disease, ageing, diabetes and cancer, will now be discussed briefly.

1.6.1 Neurological disorders: Parkinson's disease

Since the emergence of strong evidence indicating the critical role of oxidative stress in the etiopathology of PD, research has focused on the involvement of NRF2 in this disease (Cuadrado et al., 2009). Nuclear accumulation of NRF2 is seen in PD affected neurones suggesting that NRF2 activity is enhanced in the disease to increase antioxidant defences and combat oxidative stress (Ramsey et al., 2007). This interpretation is reinforced by the fact that NRF2 protects neurones from toxicity in PD mouse models using rotenone and 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP) (Todorovic et al., 2016), as well as *drosophila melanogaster* models with genetic alterations in α-synuclein (Barone et al., 2011). MPTP itself has also been shown to reduce the expression of NRF2 leading to the development of PD pathophysiology, which can be reversed by electroacupuncture which stimulates NRF2 expression (Lv et al., 2015). Also, the levels of glutathione, a major antioxidant whose synthesis is dependent on NRF2 activity, is lower in neurones from PD patients and NAC administration protects neurones against toxicity in transgenic mouse models with over expression of α-synuclein (Clark et al., 2010). Dimethyl fumarate (DMF), a potent NRF2 activator, has been shown *in vivo* to protect dopaminergic neurones against immune-related toxicity in a NRF2-dependant manner (Lastres-Becker et al., 2016; Peng et al., 2016). Also, the NRF2 activator taurodeoxycholic acid (TUDCA) has shown promising results *in vitro* (Moreira et al., 2017).

Accumulation of misfolded proteins leading to stress, is hallmark of PD. NRF2 plays a key role in modulating signalling involved in the UPR pathway and functions to stimulate α-synuclein degradation (Skibinski et al., 2017).

1.6.2 Ageing

The theory of ageing postulates that accumulative exposure the reactive oxidants over a lifetime leads to the damage of macro molecules resulting in ageing (Harman, 1956). It is now known that oxidative stress due to increase

ROS production and an impairment in antioxidant defences is a hallmark of ageing and due to its function as the master regulator of the cell oxidative stress defences it is maybe not surprising that impaired NRF2 signalling is implicated in ageing (Zhang et al., 2015). However, a lot of contradicting data has been produced, with some groups demonstrating elevated NRF2 activity in ageing and others showing decreased expression of the transcription factor, but overall it is thought that NRF2 activity declines with age (Kuosmanen et al., 2018).

Ageing studies using human bronchial epithelia cells from non-smoking young and elderly adults, show an increased basal expression of NRF2-target genes in the cells from the young adults but interestingly impaired induction of NRF2-target genes upon SFN treatment in the aged adults (Zhou et al., 2018). Also, higher ROS levels and decreased glutathione levels have been shown in the cortex of elderly patients and in NRF2 KO mice (Emir et al., 2011; Miller et al., 2012). The role that NRF2 plays in ageing has also been studied in the context of Hutchinson-Gifford progeria syndrome (HGPS), a rare premature ageing disorder found in children. HGPS is characterised by aberrant production of a mutant form of lamin protein called progerin, leading to oxidative stress. siRNA screens have shown impaired NRF2 signalling is a driver of the disease and the progerin protein itself has been shown to sequester NRF2, impairing the transcription of ARE-containing antioxidants genes, resulting in oxidative stress. Also, activation of NRF2 has been shown to reverse the phenotypic symptoms of progeria (Kubben et al., 2016). These data are further supported by work that has demonstrated that treatment with SFN enhances progerin clearance and viability in HGPS cell cultures (Gabriel et al., 2015).

1.6.3 Diabetes mellitus

Diabetes mellitus is a disease characterized by multiple pathophysiology's and high levels of oxidative stress. Activation of NRF2 through genetic alterations in *KEAP1* and pharmacologically through treatment with CDDO-Im has been shown *in vivo* to prevent the onset of diabetes (Uruno et al., 2013). Similar results have been seen with SFN treatment in a high fat fed hypoglycaemic mouse model of T2DM (Zhang et al., 2014). Diabetes patients often have impaired wound healing leading to development of chronic wound infections. Administration of RTA-408, a NRF2 -activator, has been shown to enhance wound healing in *in vivo* models of DM (Rabbani et al., 2018). Another feature of DM is impaired

vascular function which can lead to impaired cardiac function. The NRF2 activator DH404, a derivative of bardoxylone methyl, has been shown in both *in vitro* and *in vivo* models of DM to halt vascular injury.

1.6.4 Cancer

Unlike Parkinson's disease, ageing and diabetes, were the antioxidant defences are comprised due to decreased NRF2 activity, NRF2 activity is increased in cancer cells and is thought to drive tumorigenesis. Cancer cell hijack NRF2 signalling and manipulate it to their advantage by promoting angiogenesis (Zhou et al., 2012), driving proliferation (Fan et al., 2017) and enhancing chemoresistance (Wang et al., 2008a). The complex role NRF2 plays in lung cancer will be discussed in detail in Chapter 4. However, it should be noted that aberrant NRF2 signalling is not only found in lung cancer, but is also found in pancreatic (Chio et al., 2016;Lister et al., 2011), colorectal (Sadeghi et al., 2017) and breast (Zhang et al., 2016;Lu et al., 2017) cancer.

1.7 Project aims

It is well known that NRF2 plays a role in several oxidative stress related diseases and that understanding the molecular mechanisms in place to control NRF2 activity is essential to understanding the transcription factor's contribution to pathogenesis.

The aims of this thesis are:

- To establish a greater understanding of KEAP1-indepedent degradation of NRF2 through studying the involvement of a priming kinase.
- To determine the differences (if any) between mutations in *KEAP1* and mutations in *NFE2L2,* in terms of their effect on NRF2 signalling.
- To determine whether any specific vulnerabilities exist in *KEAP1* mutant cells that can be manipulated therapeutically.

2. Materials and methods

2.1 Chemicals and reagents

All chemicals were of the highest grade and were supplied by Sigma-Aldrich Ltd (Poole, Dorset, UK), unless otherwise stated. All oligonucleotides were purchased from MWG-Eurofins UK LTD (Peartree Bridge, Milton Keynes, Buckinghamshire, UK) (sequences are shown in appendix 8.1). Horseradish peroxidase (HRP)- conjugated secondary antibodies were purchased from Bio-Rad Laboratories (Perth, United Kingdom). PageRuler™ Plus Pre-stained protein Ladder, Dulbecco's Modified Eagle Medium (DMEM), Opti-MEM™, Trypsin, Lipofectamine™ 2000 Reagent, Lipofectamine RNAiMAX Reagent, Foetal calf serum (FCS), Subcloning Efficiency™ DH5α™ Competent cells and TaqMan™ gene expression assays were purchased from ThermoFisher Scientific (Paisley, Scotland, UK). Competent *Escherichia coli* (*E.Coli)* cells BL21(DE3) and Stbl3 were purchased from New England BioLabs Ltd. (Hitchin, England, UK). The following inhibitors: CT99021, MK2206, Harmine and ID-8, were all purchased from Selleckc hem (Munich, Germany). CDDO-Methyl ester (CDDO-ME) was purchased from Caymen Chemicals (Michigan, United states).

2.2 DNA cloning and analysis

2.2.1 Preparation of competent DH5α and BL21 cells

Commercial stocks of DH5α and BL21 cells were expanded to make fresh stocks of competent cells for transformation of plasmids. A 5ml inoculum overnight culture was set up using a commercial stock of competent *E.coli* cells in Luria Broth (LB) (1% w/v bacto-tryptone, 0.5% w/v bacto-yeast extract, 171 mM NaCl, H_2O). No antibiotics were added for the preparation of DH5α cells but chloramphenicol (34 μg/ml) was used for the preparation of BL21 cells. The following day, 2 ml of the overnight culture was used to inoculate 200 ml fresh LB and cultured at 37°C in a shaking incubator until the optical density (OD) at 600 nm reached 0.3-0.4. The culture was then subdivided equally into four 50 ml falcon tubes and incubated on ice for 30 min. Cells were then pelleted by centrifugation at 2,000 x*g* for 5 min at 4°C. The pellets were then resuspended in 16 ml ice cold transformation buffer 1 (10 mM $CaCl_2.2H_2O$, 10% (v/v) glycerol, 30 mM CH_3COOK (pH7.5), 50 mM $MnCl_2.4H_2O$, 100 mM RbCl -final pH 5.8 obtained using 0.2 M acetic acid) and incubated on ice for 15 min. The cells were then

pelleted again as before, followed by resuspension of the pellet in 4 ml ice cold transformation buffer 2 (75mM $CaCl_2.2H_2O$, 0.15% (v/v) glycerol, 10 mM MOPS pH 6.8, 10 mM RbCl). The cells were then stored as 100 µl aliquots at -80°C.

2.2.2 Transformation of plasmid DNA into competent cells

Competent cells (50-100 µl) were mixed with 10-100 ng DNA and incubated on ice for 30 min. A heat shock treatment of 42°C for 45 sec was given, and the reaction mix was then incubated on ice for an additional 2 min. 500 µl LB was then added, and the reaction mix was incubated at 37°C for 45 min in a shaking heat block/water bath. Cells were pelleted by centrifugation at 16,000 x*g* for 1 min, resuspended in 30 µl fresh LB (without antibiotics) and spread evenly over an LB agar (1% w/v bacto-tryptone, 0.5% w/v bacto-yeast extract, 171 mM NaCl, 1.5% w/v bacto-agar, H_2O) plate containing the appropriate antibiotic. Plates were then incubated overnight at 37°C to allow the formation of colonies which were picked the next day.

2.2.3 Plasmid DNA Isolation – plasmid miniprep and maxiprep

Bacterial colonies were inoculated in appropriate growth media (LB broth with selection antibiotics), expanded and prepped to extract plasmid DNA for sequencing and experimental use. On average three individual bacterial colonies were selected from one successfully transformed agar plate. They were screened for the plasmid of interested by inoculating a 5ml LB supplemented with antibiotic mini-culture and incubating overnight at 37°C in a shaking incubator, followed by isolation of the plasmid DNA using a QIAprep spin miniprep Kit (QIAgen, Hilden, Germany) according to manufactures guidelines, and sending to sequencing. After sequencing had confirmed the successful isolation of the desired plasmid, the construct was transformed and a 5 ml mini-culture set up as before but incubated only for 6 hr before using 4 ml of this mini-culture to inoculate an 400 ml LB-antibiotic culture. This larger culture was incubated overnight at 37°C in a shaking incubator. The next day, plasmid DNA was isolated using the HiSpeed plasmid maxi kit (QIAgen) according to manufacturer's guidelines.

The concentration and purity of isolated plasmid DNA was determined by measuring its absorbance at 260 nm using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Inc., Delaware, United States).

2.2.4 Production of glycerol stocks from bacterial cultures

Glycerol stocks were prepared for every new plasmid (either generated or resourced) for long term storage. An 800 μl aliquot of the overnight large maxiprep culture was mixed with 200 μl 50% glycerol (diluted in H_2O) in a cyrotube. Glycerol stocks were stored at -80°C. When required a small stab of the glycerol stock was removed and used to inoculate a 400 ml LB culture, supplemented with the required antibiotic.

2.2.5 Restriction digestion and ligation

To excise plasmid DNA from one vector to transfer it to another, restriction digests and ligations were carried out. All restriction digests were carried used FastDigest restriction enzymes and buffers (ThermoFisher Scientific). A 20 μl restriction digest reaction was set up containing: 5 μg DNA, 5 μl 10X FastDigest Green Buffer, 2.5 μl restriction enzyme (10,000 U/ml) and the remaining volume made up with ddH_2O. The reaction mix was incubated 2 hr in a 37°C water bath. Successful digestion of the plasmid was confirmed by running the digested plasmid on agarose gel (1% agarose in TAE (tris base, glacial acetic acid, 0.5 M EDTA pH 8, H_2O)) and comparing it with a non-digested control. After digestion, the insert DNA was excised from the agarose gel and purified. Purified insert DNA was then ligated with new backbone vector digested with the same restriction enzymes. Ligation was carried out overnight at room temperature using 1 μl of 1 U/μl T4 DNA ligase enzyme (ThermoFisher Scientific) using vector and insert DNA in 1:3 molar proportion.

2.2.6 Gel extraction

To isolate a cut fragment of DNA after restriction enzyme digestion of a plasmid, gel extraction was carried out post-electrophoresis. The desired cut fragment of DNA was carefully excised from the gel using a scalpel. The fragment was purified using the QIAquick gel extraction kit (QIAgen) according to manufacturer's guidelines.

2.2.7 Site directed mutagenesis

Primer pairs to insert a single amino acid substitution were designed to include the minimal number of base substitutions to encode the new desired amino acid. All primer sequences are listed in Appendix 8.1. The thermal cycling conditions

for the polymerase chain reaction (PCR) reaction are as described below, and the reaction set up was as described in Table 2.2.1.

1. Initial denaturation at 95°C for 2 min
2. Denaturation at 95°C for 20 sec
3. Annealing at 60°C for 10 sec
4. Extension at 68°C for 30 sec per Kb plasmid length
5. Final extension at 68°C for 5 min

} 20 cycles

Following PCR, DNA was further digested with 1.5 μl Dpn1 (per reaction) for 1 hr at 37°C to ensure full digestion of any potential methylated parental DNA template. The digested product was then transformed and plated onto specific antibiotic containing agar plates (see section 2.2.2 for details).

Component and stock concentration	Volume required for 1 reaction (30 μl)
Template DNA	100 ng
Forward primer (100 μM)	1.5 μl
Reverse primer (100 μM)	1.5 μl
10X PCR Buffer for KOD Hot Start DNA Polymerase	5 μl
2 mM dNTPs	5 μl
2.5 mM $MgSO_4$	5 μl
200 U KOD Hot start DNA Polymerase	1 μl
dH_2O	X μl

Table 2.2.1 : PCR reaction set up for site directed mutagenesis

2.2.8 DNA sequencing

Sequencing was carried out after site-directed mutagenesis and ligation reactions to confirm the correct nucleotide sequence using the sequencing primers listed in Table 2.2.2. The Genetic Core Services at Ninewells Hospital and Medical School, carried out all DNA sequencing.

Vector/Insert	Primer name	Sequence (5-3')
pGEX	pGEX Forward	ATA GCA TGG CCT TTG CAG G
pcDNA 3.1- V5/His	T7	TAA TAC GAC TCA CTA TAG GG
LentiCRISPRV2	SeqLentiCRISPRV2 HR	GAG GGC CTA TTT CCC ATG ATT
Δ17-32 mNrf2 pcDNA 3.1 v5/His	Δ17-32 mNrf2 Internal 1 Forward	GAA GGC CTT TTT TGC TCA GTT TCA AC
	Δ17-32 mNrf2 Internal 2 Reverse	AGA ATC CTC AAA ACC ATG AAG GAA ATG TG
	Δ17-32 mNrf2 Internal 3 Forward	CAC ATT TCC TTC ATG GTT TTG AGG ATT CTT TC

	Δ17-32 mNrf2 Internal 4 Forward	CGG AAA TGG AGG AGC TAG ATA GTG CCC CTG G
GST-mNeh6^{+} pGEX / mNeh6- LacZ	mNeh6^{+} Forward	GCC TTC CTC CGC TGC CAT CAG TCA GTC ACT CTC TG
	mNeh6^{+} Reverse	CAC ATT GGG ATT CAC GCA TAG GAG CAC TGT TC
	pGEX Forward	ATA GCA TGG CCT TTG CAG G
GST-hNeh6^{+} pGEX	pGEX Forward	ATA GCA TGG CCT TTG CAG G

Table 2.2.2: Details of primers for DNA sequencing

2.2.9 Generation of KEAP1KO CRISPR plasmids for stable cell line production

To generate a panel of isogenic cell lines with multiple different genetic alterations, several DNA constructs were generated in the LentiCRISPRV2 vector. The LentiCRISPRV2 plasmid and guide sequence for the generation of NRF2 knockout (KO) and NRF2 gain-of-function (GOF) cells were kindly gifted by Dr Laureano de la Vega (University of Dundee). This guide sequence was directed against the DLG and ETGE KEAP1 binding regions found in the Neh2 domain of NRF2. Clones that do not repair in frame lead to generation of a NRF2 KO and those that repair in frame through alteration in the KEAP1 binding site generate a NRF2 GOF mutant. For the generation of KEAP1 KO cell lines paired oligonucleotide primers were designed against a region early on in exon 1, to ensure the majority of the KEAP1 protein is not transcribed. All guide sequences are provided in Appendix 8.2.

Firstly, 5 μg LentiCRISPRV2 plasmid was digested and dephosphorylated with *BsmBI* for 30 min at 37°C. The composition of this reaction mix was as described in Table 2.2.3.

Component and stock concentration	**Volume/ amount required (total volume 60 μl)**
LentiCRISPRV2	5 μg
FastDigest *BsmBI*	3 μl
FastAP (1U/μl)	3 μl
10X FastDigest Buffer	6 μl
100 mM DTT (freshly prepared)	0.6 μl
ddH_2O	34.7 μl

Table 2.2.3 : Reaction composition for digestion and dephosphorylation of LentiCRISPRV2 plasmids

The DNA was then ran on a 1% agarose gel in TAE buffer. The larger (~15 Kb) band of the digested vector plasmid was then gel purified using the QIAquick gel extraction kit. Each pair of oligonucleotides were then phosphorylated and annealed together as described in Table 2.2.4 under the following thermocycler conditions:

1. 37°C for 30 min
2. 95°C for 5 min then ramped down to 25°C at 5°C/ min

Component and stock concentration	Volume required (total volume 10 μl)
Oligonucleotide 1 (100 μM)	1 μl
Oligonucleotide 2 (100 μM)	1 μl
10X T4 ligation buffer (New England BioLabs)	1 μl
T4 polynucleotide kinase (New England BioLabs)	0.5 μl
ddH_2O	6.5 μl

Table 2.2.4: Reaction set up for phosphorylation and annealing of oligonucleotides

The annealed oligonucleotide pairs were then diluted 1:200 using ddH_2O. A ligation reaction to inset the oligonucleotides into the LentiCRISPRV2 backbone was set up as described below in Table 2.2.5.

Component and stock concentration	Volume/ amount required (total volume 11 μl)
BsmBI digested plasmid	50 ng
Diluted oligonucleotide pairs	1 μl
2X Quick ligase buffer (New England BioLabs)	5 μl
Quick ligase (New England BioLabs)	1 μl
ddH_2O	3 μl

Table 2.2.5: Ligation reaction set up for generation of LentiCRISPRV2 plasmids

The whole ligation reaction mix was then transformed into Stbl3 competent cells and spread evenly across LB agar plates supplemented with ampicillin (see section 2.2.2). Colonies were selected followed by DNA preparation using QIAprep spin miniprep kit (see section 2.2.3) and sent of sequencing (see section 2.2.8). After correct insertion was confirmed through sequencing, larger quantities of DNA were prepared using HiSpeed plasmid maxi kit and stored at -20°C until use. Also, glycerol stocks were generated for long term storage (see section 2.2.4).

2.2.10 Genomic DNA extraction and PCR reaction

To determine the exact location of the genetic alteration and whether the mutations generated are homozygous or heterozygous, genomic DNA was extracted from the CRISPR cell lines and sent for sequencing. Cells were allowed to reach a confluency of 90% in a 6-well plate before harvesting genomic DNA. Cells were trypsinized, pelleted and then genomic DNA was extracted according to the manufactures protocol provided with the Nucleospin® Tissue kit (Macherey-Nagel, Duren, Germany).

2.2.11 Sequencing of CRISPR cell lines

Sequencing of the CRISPR cell lines was kindly carried out by Mr Thomas McCartney (MRC Protein Phosphorylation and Ubiquitination centre, University of Dundee). Shotgun cloning of the pooled PCR product (generated in 2.2.10) into TOPO vectors was carried out. Several clones per cell line were then selected, restriction digested with *EcoRI* to confirm the insert and then sequenced with a M13F primer (sequence 5-3': GTA AAA CGA CGG CCA GT).

2.3 Protein purification and analysis

2.3.1 Recombinant protein expression in *E.Coli*

To generate purified proteins for use in *in vitro* kinase assays, plasmid DNA was expressed in BL21 cells. All glutathione S-transferase (GST) fusion proteins were expressed and purified using *E.Coli* strain BL21 cells. 50 ng sample of the desired plasmid was transformed into BL21 cells and plated on LB-agar plates containing 100 μg/ml ampicillin and 34 μg/ml chloramphenicol. Plates were incubated at 37°C overnight in an incubator. The following day 2-4 single colonies were picked from the plates and grown overnight in a 10 ml culture of LB selection media, to prepare a starter culture.

400 ml LB selection media was then inoculated with 4 ml above overnight starter culture and grown at 37° C in a shaking incubator. The OD of the bacterial culture was routinely monitored until an OD at 600 nm of 0.4-0.6 was reached. A 1 ml aliquot of the culture was collected at this point as a pre-induction sample. 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to the remaining culture and the cells were allowed to grow at 30°C under shaking conditions for 4 hr. Another 1 ml aliquot was collected as a post-induction sample. The cells were pelted by centrifugation at 6,000 x*g* for 15 min at 4°C, the supernatant was

discarded and the pellet stored at -80°C. To confirm successful protein induction, both pre- and post-induction samples were pelleted by centrifugation at 13,000 x*g* for 1 min and lysed in 30 μl of 2X NuPAGE™ LDS sample buffer (ThermoFisher Scientific), before running the samples on a 10% sodium dodecyl sulfate (SDS)-polyacrylamide gel (SDS-PAGE). The gel was stained with 5-10 ml SimplyBlue™ SafeStain coomassie (ThermoFisher Scientific) until the appearance of visibly clear distinguishable bands (approximately 1 hr). Coomasssie was then removed and the gel washed three times for 5 min intervals with water to reduce the background stain. Expression of the desired recombinant protein was confirmed by bands being present at the correct molecular weight in the post- induction sample and absent from the pre- induction sample.

2.3.2 Protein purification

All proteins expressed in BL21 cells were purified before experimental use. After confirming IPTG induced protein expression, the cell pellet was lysed in 10-15 ml B-PER™ complete bacterial protein extraction reagent (ThermoFisher Scientific) depending on the size of the pellet. The pellet was vortexed and the 15 ml lysate split equally into two 15 ml falcon tubes before placing on an orbital rotor at 4°C for 30 min. Following lysis, the lysate was centrifuged at 10,000 x*g* for 20 min at 4°C and the supernatant collected. The supernatant was then passed through 0.22 μm sterile filter and combined in one 15 ml falcon tube. GST-fusion proteins were all affinity-purified using Glutathione Sepharose® 4B beads (Sigma-Aldrich), the beads were prepared by repeatedly washing the beads with 10 volumes of phosphate buffer saline (PBS). 1 ml of GST beads was then incubated with 15 ml of lysate for 2 hr on an orbital rotator at 4°C. The beads were then pelleted by centrifugation at 13,000 x*g* for 1 min, and the supernatant retained, then washed three times for 5 min intervals with 10 ml wash buffer (50 mM tris pH 7.5, 150 mM NaCl) on a rotor. The GST-fusion protein was then eluted from the beads using 500 μl of elution buffer (50 mM tris pH 8.8, 150 mM NaCl, 10 mM reduced glutathione, final pH adjusted to pH 7.5) at room temperature for 30 min. The elution step was repeated twice and 20 μl of each elution was lysed directly in 2X NuPAGE™ LDS sample buffer, before running on a 10% SDS-polyacrylamide gel. Successful purification was confirmed by the presence of clear distinguishable bands at the correct molecular weight for the GST-fusion protein. Eluted proteins were then combined and injected into a Slide-A-Lyzer™

dialysis cassette (10 K MWCO, ThermoFisher Scientific) then dialysed overnight at 4°C in dialysis buffer (50 mM tris pH 7.5, 150 mM NaCl, 10% glycerol), to remove the reduced glutathione. The concentration of the purified protein was then estimated by carrying out a detergent compatible (DC) assay, followed by SDS- PAGE and staining with SimplyBlue™ SafeStain coomassie using known bovine serum albumin (BSA) standards. Proteins were aliquoted and stored at -80°C.

2.4 Radioactive *assays*

2.4.1 *In vitro* kinase MBP activity assays

Before determining the ability of various kinases to phosphorylate NRF2 mini-proteins, activity assays were carried out with each kinase against a well characterised substrate. All kinases, except for DYRK4 which was procured from ThermoFisher Scientific, were kindly provided by Dr James Hastie (Division of Signal Transduction Therapy, University of Dundee). Upon receiving, all kinases were assayed against Myelin Basic Protein (MBP). These activity assays ensured the same number of units of kinase were added to each reaction so the ability of the kinase to phosphorylate NRF2 could be compared with the ability of the kinase to phosphorylate MBP, a well-known substrate that is heavily phosphorylated. The details of the kinases the assay set up are provided in Table 2.4.1.

Component	Volume required in a 50 μl assay
0.1 μg Kinase	5 μl
MBP (1 mg/μl)	15 μl
Kinase Assay Buffer (50 mM Tris-Cl, 0.03% (v/v) Brij-35, 0.1% (v/v) β-mercaptoethanol)	25 μl
10 mM $MgCl_2$ and 0.1 mM [γ-^{32}P]-ATP (approx. 0.5x10^6 cpm/nmole)	5 μl

Table 2.4.1: Reaction set up for kinase activity assays

Kinases were diluted to 0.1 μg with kinase dilution buffer (50 mM Tris-Cl, 0.03% (v/v) Brij-35), then mixed with MBP and kinase assay buffer. The reaction start times were staggered by 15 sec to ensure that each reaction was incubated for the same time period, by pipetting the 10 mM $MgCl_2$ and 0.1 mM [γ-^{32}P]-ATP onto the side on the tube and then flicking it down into the reaction mixture and placing the tubes in a 30°C heat block while simultaneously starting the timer. The reactions were incubated at 30°C for various time points to ensure initial rate

conditions were met, before 40 μl of each reaction was spotted onto pencil numbered Whatman™Grade P81 ion exchange cellulose chromatography paper (Fisher Scientific, product has been discontinued) (1.5x1.5cm^2). The papers were then dropped into 75 mM $H_3(PO_4)$ to remove any unreacted ATP and then washed four times with 75 mM $H_3(PO_4)$ for 5 min intervals, using a magnetic stirrer. For the final wash the papers were placed in acetone, and then air dried prior to scintillation counting in 2 ml Optifluid. As a control two "blanks" were set up in parallel that had no kinase present in the reaction mixture, the counts for these tubes were then subtracted from all experimental counts. For every new radioactive ATP stock pot used, a new ATP counter was made by diluting a 10 mM ATP probe 1:100 with kinase dilution buffer and taking 10 μl of the dilution into a scintillation tube with 2 ml Optifluid. The ATP counter decays at the same rate as the ATP stock used in the experiment and therefore helped to determine the activity of the kinase.

The kinase activity was calculated using the following formula:

*Kinase activity (*pmol/min/μl) = *(((scintillation counts – blank counts)/ (ATP activity) / (Assay))) time*.

2.4.2 *In vitro* kinase assays

In vitro kinase assays were carried out to determine whether a kinase may target several different serine residues in the Neh6 domain of NRF2. Purified recombinant GST-tagged mNeh6$^+$ protein, GST-tagged hNeh6$^+$ or single amino acid mutants of GST-tagged m/hNeh6$^+$ proteins, generated as described in section 2.2.7, were incubated with purified activated recombinant kinase in the presence of 10 mM $MgCl_2$ and 0.1 mM [γ-^{32}P]-ATP in a 30°C heat block, composition of the reaction mix is given below in Table 2.4.2. The length of the assay varied between the four groups of kinases, dual specificity tyrosine-phosphorylation regulated kinase (DYRK) and stress-activated protein kinase (SAPK) family members were subjected to a 10 min assay whereas JNKs and the "miscellaneous" kinases ((mixed-linage kinase-1 (MLK1), thr/tyr kinase (TTK), TANK-binding kinase 1 (TBK1) and cyclin-dependent kinase 9 (CDK9)) were tested in a longer 30 min assay. These time points were decided from the previous activity assays described above in section 2.4.1. All reactions were terminated by the addition of pre-warmed (heated at 75°C for 10 min) 4X LDS gel

loading buffer followed by boiling the sample for 10 min, the sample were then loaded on to a SDS-PAGE gel, and the excess [γ-^{32}P]-ATP was removed by electrophoresis. Gels were then stained with SimplyBlue™ SafeStain coomassie, dried onto filter paper and subjected to autoradiography. After which, the protein bands were excised and the incorporated radioactivity was determined by scintillation counting. “Blanks” containing only the purified protein but no kinase were included. Scintillation counts for blank reactions were subtracted from experimental readings to account for background radiation. The kinase activity towards the Neh6^{+} was then calculated using the formula given in section 2.4.1.

Component	Volume/ concentration required for 40 µl reaction
Purified protein (GST-mNeh6^{+}, GST-hNeh6^{+} or single amino acid mutants)	4 µg
Kinase	2 mU/µl
Kinase assay buffer (50 mM Tris-Cl, 0.03% (v/v) Brij-35, 0.1% (v/v) β-mercaptoethanol)	20 µl
10 mM $MgCl_2$ and 0.1 mM [γ-^{32}P]-ATP (approx. 0.5x10^{6})	4 µl

Table 2.4.2: Reaction set up for in vitro kinase assays

2.4.3 *In vitro* two-step linked kinase assay

To determine whether a kinase works in concert with GSK-3β to enhance GSK-3β mediated phosphorylation of NRF2, *in vitro* two-step linked kinase assays were carried out.

In the first step, purified recombinant GST-mNeh6^{+} protein (0.1 µg/µl, substrate) or single base pair mutant GST-mNeh6^{+} proteins (0.1 µg/µl, substrate) generated from BL21 cells, were incubated with 2 mU/µl purified activated recombinant DYRK2, in the presence of 5 mM $MgCl_2$ and 0.05 mM cold- [γ-^{32}P]-ATP in a 30°C heat block for 2 hr. The reaction mix was then heated at 60°C for 5 min to inactivate the kinase present, and then re-incubated at 30°C for 10 min to bring the reaction mix down to ambient temperature. Purified activated recombinant GSK-3β (6 mU/µl) along with of 5 mM $MgCl_2$ and 0.05 mM [γ-^{32}P]-ATP was then added and the reaction incubated at 30°C for a further 1 hr. All reactions were terminated by the addition of hot (heated at 75°C for 10 min) 4X LDS gel loading buffer followed by boiling the samples for 10 min, the samples were then loaded on to an SDS-PAGE gel, and the excess [γ-^{32}P]-ATP was removed by electrophoresis. Gels were then stained with SimplyBlue™ SafeStain

coomassie, dried onto filter paper and then subjected to autoradiography. The protein bands were excised and the incorporated radioactivity was determined by scintillation counting. “Blanks” containing only the purified protein but no kinase were included. Scintillation counts for blank reactions were subtracted from experimental readings to account for background radiation. The kinase activity towards the Neh6 $^{+}$ was then calculated using the formula given in section 2.4.1.

2.5 Isolation, quantification and analysis of cellular protein

2.5.1 Protein extraction

For the preparation of protein lysates for western blotting, protein was first extracted from cells using the following protocol. Cells were seeded and treated according to the requirements for the experiment. The media was aspirated off the cells and the cells were washed twice in PBS and then incubated in kinase lysis buffer (10%(w/v) sucrose, 50 mM tris-HCL pH 7.5, 1 mM EDTA pH 8.0, 1 mM EGTA pH 8.0, 2 mM sodium vanadate (Na_2VO_3), 10 mM glycerophosphate, 25 mM sodium fluoride (NaF), 5 mM sodium pyrophosphate (NaP_2O_5), 1% (v/v) NP-40) supplemented with cOmplete™ EDTA- free protease inhibitor cocktail tablets (Sigma-Aldrich) for 15 min on ice. Cells were then manually scraped into the buffer, transferred to an eppendorf and vortexed briefly before centrifugation at 12,000 x*g* for 10 min at 4°C.

2.5.2 Protein quantification

To ensure equal loading of protein samples for western blotting, protein quantification and normalisation was carried out on all new lysates. Total protein quantification was performed on the clarified lysates using a Bio-Rad™ DC protein assay kit (Bio-Rad) in a 96-well plate format. 5 μl of sample lysate was incubated with a 25 μl of a 50:1 (v/v) ratio of combined reagent A and reagent S, in triplicate in a 96-well plate. 200 μl of reagent B was then added to each well, before incubating the plate in the dark at room temperature for 15 min. The absorbance of each sample was then measured at 750 nm using a SpectraMax M2 (Molecular Devices, California, United States).

The concentration of the protein in the cell lysate (mg/ml) was then quantified against BSA standards of known concentrations (4, 2, 1, 0.5, 0.25, 0.125, 0.063 and 0 mg/ml). Samples were then diluted to a final concertation of 1 μg/μl with a 1:10 mixture of kinase lysis buffer and 4X NuPAGE ™ LDS sample

buffer. Samples were then vortexed thoroughly, followed by heat denaturation of proteins by boiling at 100°C for 10 min, before loading onto a SDS-PAGE gel.

2.5.3 SDS polyacrylamide gel electrophoresis using handmade gels

Proteins in samples were separated according to their molecular weight through running down a SDS-PAGE gel. Denatured protein samples (typically 10-20 μg) were loaded onto a SDS-PAGE of the desired density, along with 4 μl PageRuler™ Plus Pre-stained protein ladder and separated according to the molecular weight of the protein. The gels were set in two layers; a bottom resolving gel and a top stacking gel, the composition of both gel layers is shown in Table 2.5.1.

The gels were run using ThermoFisher Scientific mini gel tank at 100 V in running buffer (25 mM tris, 250 mM glycine, 0.1% (w/v) SDS) until the samples had successfully passed through the stacking gel. The voltage was then increase to 190 V and ran until the dye front had almost reached the bottom of the gel.

Component	Resolving gel	Stacking gel
30% 37.5:1 Acrylamide:Bis solution (National diagnostics)	6-12% (v/v)	5% (v/v)
1.25 M tris pH 6.8	-	126 mM
1.25 M tris pH 8.8	375 mM	-
10% (w/v) SDS	0.1% (w/v)	0.1% (w/v)
10% ammonium persulfate (APS)	0.1% (w/v)	0.1% (w/v)
N,N,N',N'-Tetramethylethylenediamine (TEMED) (Sigma-Aldrich)	0.04% (w/v)	0.01% (w/v)

Table 2.5.1: Composition of stacking and resolving gels for SDS-PAGE

2.5.4 SDS poly-acrylamide gel electrophoresis using pre-cast gels

The electrophoresis of radioactive experiments was carried out using pre-cast gels. Novex NuPAGE™ 4-12% Bis-tris pre-cast protein gels were run using ThermoFisher Scientific mini gel tanks and 1x NuPage™ MOPS SDS running buffer at 150V, until the dye front nearly reached the end of the gel. The gel was either transferred for western blotting or stained with SimplyBlue™ SafeStain coomassie for protein quantification.

2.5.5 Western blotting

To allow the visualisation of proteins, gels were transferred to membranes, incubated with antibody and imaged, according to the following. After SDS-PAGE, resolved proteins were then transferred onto either polyvinyl difluroride (PVDF) (ThermoFisher Scientific) or nitrocellulose membranes (ThermoFisher Scientific). All PVDF membranes were pre-soaked in 100% methanol immediately prior to use. The Bio-Rad Trans-Blot ® cell system was used to transfer proteins to the membrane at 100V for 60 min using transfer buffer (25 mM tris, 250 mM glycine, 20% (v/v) methanol) in the presence of an ice pack. The transferred membrane was then blocked in 5% (w/v) dried skimmed milk (Marvil) in TBS-T (50 mM Tris, 150 mM NaCl, 0.25% (v/v) Tween 20®) for 1 hr at room temperature with agitation. Post-blocking, the membrane was incubated for 3-16 hr in the appropriate primary antibody diluted in 5% dried skimmed milk-TBS-T (Appendix 8.3 antibodies and dilutions). The membranes were then washed 3 times with TBS-T for 5 min each time before incubation with the required HRP-conjugated secondary antibody diluted 1:2,000 or near-infrared fluorescent secondary antibody at a dilution of 1: 15,000 (LI-COR Biosciences, Nebraska, United states) in TBS-T containing 5% (w/v) dried skimmed milk for 1 hr at room temperature. The membrane was the again washed a further 3 times with TBS-T for 5 min each. Membranes incubated with LI-COR secondary antibodies were shielded from light throughout the incubation and subsequent washing stages.

For the imagining of membranes incubated with HRP-secondary antibodies, enhanced chemiluminesence (ECL) and autoradiography film were used to visualise the antibody-protein complexes. After the final wash the membranes were incubated with 1:1 mix of in-house ECL solution 1 (100 mM tris pH 8.5, 2.5 mM luminol, 0.4 mM ρ-coumaric acid) and ECL solution 2 (100 mM tris pH 8.5, 0.02% (v/v) hydrogen peroxide) before visualisation using autoradiography X-ray film. Membranes probed with near-infrared fluorescent secondary antibodies were imaged using a LI-COR Odyssey ® CLx scanner. Quantification against a house keeping protein was carried out using the in-built LI-COR Image Studio software.

2.6 Cell culture and cell-based assays

2.6.1 Cell culture conditions

Cell lines described the table below, were all cultured in a humidified 5%CO_2 incubator at 37°C. All cell lines were maintained in DMEM supplemented with 10% (v/v) FCS. Table 2.6.1 lists all the cell lines used in this thesis along with their original source.

Cell line	Description	Base medium	Origin
A549	Human alveolar basal adenocarcinoma epithelial cells	DMEM	Hayes lab stocks
NRF2 KO A549	Genetically engineered CRISPR cells lines where the gene encoding NRF2 has been knocked out	DMEM	Dr Laureano de la Vega
H460	Human lung pleural effusion cell line	DMEM	Hayes lab stocks
H1299	Human non-small cell lung adenocarcinoma cell line	DMEM	Dr Laureano de la Vega
NRF2 KO H1299	Genetically engineered CRISPR cells lines were the gene encoding NRF2 has been knocked out	DMEM	Generated in lab
NRF2 GOF H1299	Genetically engineered CRISPR cell lines were the gene encoding NRF2 has been altered	DMEM	Generated in lab
KEAP1 KO H1299	Genetically engineered CRISPR cell lines were the gene encoding KEAP1 has been knocked out	DMEM	Generated in lab
H2228	Human non-small cell lung adenocarcinoma cell line	DMEM	ATCC
H1568	Human lung adenocarcinoma cell line derived from a metastatic site; lymph node	DMEM	ATCC
H1395	Human lung adenocarcinoma epithelial cells	DMEM	ATCC
DLD1	Human colorectal adenocarcinoma epithelial cells	DMEM	Dr Laureano de la Vega

Table 2.6.1:Details of cell lines and culture conditions

2.6.2 Sub-culturing cells

Cell lines were sub-cultured every two to three days or when they reached a confluency of 80-90%. Cell growth media was removed and the cells were washed twice with pre-warmed sterile 1X PBS before the addition of 2-3 ml of 0.025% trypsin (ThermoFisher Scientific) supplemented with 0.05% (v/v) EDTA

(0.05 mM). Upon addition of trypsin the flasks were re-incubated for 2-3 min at 37°C, after which, any un-detached cells were cleared from the back wall through gentle agitation of the flask. Detached cells were re-suspended in fresh media, and the desired number of cells was transferred into a new flask containing fresh media and returned to the incubator.

2.6.3 Cryopreservation of cells

Any newly generated or newly obtained cell line in the lab was expanded and frozen down for long term storage in liquid nitrogen. Cells from a confluent flask were trypsinized and re-suspended in DMEM supplemented with 10% FCS to a final volume of 10 mls. Cells were then pelleted by centrifugation at 12,000 x*g* for 3 min. The growth media was then removed and the cell pellet was re-suspended in freezing media (90% FCS and 10% DMSO). Cells were frozen in 1ml aliquots (approximately four 1 ml aliquots were obtained from a confluent T75-flask) in individual cryogenic vials (Corning, New York, United states). The cyrovials were stored at -80°C in a cryofreezing container overnight, before they were transferred to liquid nitrogen. When required, cells were revived from liquid nitrogen storage, by thawing the cells rapidly in a 37°C water bath. The thawed cells were then transferred in to a T-25 flask containing pre- media and incubated overnight. The next day, the cells were washed twice with sterile 1X PBS and the media was renewed. Cells were only used in experiments after they had undergone three passages after liquid nitrogen storage and had tested negative for mycoplasma contamination.

2.6.4 Seeding cells

Cells were counted using a haemocytometer and seeded at different densities depending on the experimental requirements. Cells were trypsinized and re-suspended thoroughly in a final volume of 10 ml of media. A volume of 20 μl of the cell suspension was then mixed with 20 μl 0.4% trypan blue solution (Thermo Fisher Scientific) and 10 μl of the mix was counted with a Countess II FL automated cell counter (ThermoFisher Scientific). The cell suspension was then diluted with media to obtain the required cell count for the experiment.

2.6.5 Transient transfection using Lipofectamine 2000™

Transient transfections were carried out for the incorporation of plasmid DNA into mammalian cells**.** Cells were seeded in 6-well plates 24 hr prior to

transfection to achieve a confluency of approximately 80% the following day. DNA (1-2 μg) and Lipofectamine 2000™ (μl) were transfected at a ratio of 1:3 into the desired cell line. Firstly, the plasmid DNA was combined with OptiMEM™ and allowed to incubate for 20 min at room temperature. In a separate tube, Lipofectamine 2000™ was combined with the same volume of OptiMEM™ as the plasmid. Both transfection mixes were then combined and incubated for a further 20 min, before gently adding drop wise to the wells containing 1.8 ml of fresh media. Cells were then either harvested 24 hr post-transfection or treated with drugs to be harvested at a later time point as per the experimental plan.

2.6.6 Transient knock-down siRNA reverse transfections using Lipofectamine™ RNAiMAX

Short interfering RNA (siRNA) was transiently transfected into mammalian cells to knockdown genes of interest. A volume of 500 µl OptiMEM™ was combined with 20 nM siRNA (negative control or specific against the gene of interest) in a tube and mixed well. This was then transferred to one well of a 6-well plate and the side of the plate agitated to ensure complete coverage of the well with the mixture. 5 µl of Lipofectamine ® RNAiMAX was then added to the well and incubated for 10 min at room temperature. Cells were trypsinzed and counted, according to sections 2.6.2 and 2.6.4 above, to plate $2x10^5$ cells per 2 ml/well in a 6-well plate. Cells were incubated for 48 hr before confirming the magnitude of the target gene knockdown at both a protein (through western blotting) and an mRNA level (through qPCR). Both the KRAS SMART pool of ON-TARGET plus siRNA and the negative control siRNA were purchased from Dharmacon (Dharmacon, Colorado, United States).

2.6.7 Transfection for the generation of stable CRISPR/Cas9 cell lines

For the generation of stable isogenic Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) cell lines with genetic alterations, plasmid DNA generated as described in section 2.2.8 was transfected into H1299 cells. H1299 cells were seeded in 6-well plates 24 hr prior to transfection to achieve a ~80% confluency the following day. 2 µg of DNA for one of the following constructs was transfected into the cells as described in section 2.6.5: LentiCRISPRV2 empty plasmid, KEAP1KO-LentiCRISPRV2 or NRF2KO-LentiCRISPRV2. One well was transfected with 2 µg of pMAX-GFP as a control to check transfection efficiency.

At 48 hr post-transfection the cells were treated with 2 µg/ml puromycin (Puromycin dihydrochloride; ThermoFisher Scientific) to kill any cells not possessing the resistance cassette from the LentiCRISPRV2 plasmid. A control well of H1299 cells that had not undergone any transfection was included, this well was treated with puromycin at the same time. Once all the cells in the control well had died, the cells in the remaining transfected wells were then trypsinized and transferred to a T-25 flask, this is the "pool" population. These flasks were then expanded, harvested for protein and RNA analysis, and then frozen liquid nitrogen cell stocks were produced.

2.6.8 Colony selection for CRISPR cell lines

To produce single cell clones the cells from the "pool" population were counted (as described in section 2.6.4) and seeded in 96-well plates. Three seeding densities were chosen: 0.5 cells per well, 1 cell per well, and 2 cells per well. Triplicate 96-well plates were seeded per density and each plate was screened regularly by eye to check for the presence of single colonies. Wells containing single colonies were then trypsinized and expanded into 24-well plates. Once the wells became confluent they were then further expanded into 12-well plates and samples taken for western blotting and qPCR analysis to confirm genetic alteration. Positive clones where then expanded and liquid nitrogen stocks were produced.

2.6.9 ARE luciferase reporter assay

To determine the effect that phosphorylation of several serine resides in the Neh6 region of NRF2 has on NRF2-activity, luciferase reporter assays were carried out. NRF2 KO A549 cells were seeded in 6-well plates and co-transfected with an empty pcDNA3.1-V5/His expression vector or various ∆17-32 mNrf2 pcDNA3.1-V5/His single amino acid site mutant expression plasmids (generated as described in section 2.2.7) along with the murine quinone reductase-based P-1016/nqo1-Luc reporter construct and pRL-TK Renilla. A volume of 2 µg of each plasmid was transfected into the cells in each different experimental group. Cells were transfected as described in section 2.6.6. 24 hr post-transfection the cells were washed and the media replaced with 0.1% FCS DMEM to serum starve the cells. Cells were serum starved overnight for 16 hr before treatment with inhibitors; 5 µm MK2206, 5 µM CT99021,10 µM Harmine or a DMSO control for a further 18 hr in reduced serum 0.1% FBS DMEM. Plates were then washed

twice with PBS then harvested in 1X passive lysis buffer and transferred into white-walled clear bottom 96-well plates (Corning). The assay was carried out according to Promega guidelines for the Dual-Luciferase® Reporter Assay System (Promega, Wisconsin, United states) using a GloMax®-Multi^{+} Detection System with two injectors (Promega). All results were the normalised against the Renilla luciferase activity to account for transfection variability between experimental groups.

2.6.10. β-galactosidase assay

To determine what effect phosphorylation of several serine residues in the Neh6 region has on NRF2 protein stability, β-galactosidase assays were carried out. NRF2 KO A549 cells were seeded in 6-well plates and co-transfected with pcDNA3.1 expression vector for a V5-tagged fusion protein comprising mNeh6^{+} coupled C-terminally to LacZ, this plasmid is referred to as mNeh6^{+}LacZ-V5, or with single amino acid site mutants of the same plasmid (generated as described in section 2.2.7). An empty LacZ-V5 plasmid was included as a control. At 24 hr post-transfection the cells were serum starved by removing the media and replacing it with 0.1% FBS DMEM. After 16 hr of serum starvation, the cells were then drugged for a further 18 hr with inhibitors: 5 μm MK2206, 5 μM CT99021, 10 μM Harmine or a DMSO control, also in reduced serum conditions. The plates were then washed twice in PBS and the cells harvested in 1x reporter lysis buffer and transferred into white-walled clear bottom 96-well plates. β-galactosidase activity was then measured according to the Promega Guidelines given in the β-galactosidase Enzyme Assay System with Reporter Lysis Buffer kit (Promega) using a GloMax-Multi+ Detection System with two injectors. All results were normalised against total protein concentration, by harvesting the cells after the assay, according to the description given in sections 2.5.1 and 2.5.2.

2.6.11 Fluorescent measurement of glutathione using mBC

Monochloromobimane (mBC) is a non-fluorescent probe that when taken up into cells reacts with low molecular weight thiols, such as glutathione, and fluoresces. This resulting florescence can then be correlated to the amount of glutathione in the cell. Cells were seeded in 6-well plates so that after 24 hr they had reached ~80% confluency. The wells were washed twice with Dulbecco's phosphate buffered saline (DPBS, ThermoFisher Scientific) with no calcium, magnesium or phenol red, and 1 ml mCB (ThermoFisher Scientific) diluted in

hanks balanced salt solution (HBSS, ThermoFisher Scientific) to a final concentration of 20 μM was added to the wells. HBSS alone was added to one well as a control and the value for this control well was taken away from all experimental readings to account for background cell florescence. Plates were then incubated in the dark for 30 min at room temperature. Fluorescence was measured using a SpectraMax M2 plate reader, with excitation at 390 nm and emission at 490 nm. All results were normalised against total protein concentration, by harvesting the cells after the assay, following the descriptions given in sections 2.5.1and 2.5.2.

2.6.12 Determination of total, oxidized and reduced glutathione from cells using the enzymatic recycling method

For a more robust measurement of glutathione and to obtain specific values for total, oxidized and reduced forms a protocol modified from Rahman et al. (2007) was used. Cells were seeded in 6-well plates so that on the day of the assay they had reached a confluency of ~80%. At 24 hr post-seeding cells were drugged with either 100 mM buthionine-sulfoximine (BSO) (Sigma-Aldrich) for 24 hr or 5 μM NAC (Sigma-Aldrich) for 3 hr. Media was aspirated from plates and wells were washed twice with room temperature PBS. Cells were then manually scrapped from the back wall of the plate into 1 ml of PBS, followed by pelleting via centrifugation at 13,000 x*g* for 5 min at 4°C. Supernatant was carefully aspirated and pellets were flash frozen (either on dry ice or by submerging in liquid nitrogen). Pellets were re-suspended in 250 µl 0.25 M sucrose (sucrose dissolved in distilled water) and a 50 µl aliquot was taken for protein quantification according to section 2.5.2. A volume of 200 µl 2X extraction buffer (0.2% v/v Triton X-100, 0.2% v/v NP-40, 1.2% m/v 5-sulfosalicylic acid, 10% meta-phosphoric acid, made up to 20 ml with 2x KPE (0.2 M potassium phosphate, 10 mM EDTA)) was added to the remaining 200 µl of sample and mixed through briefly vortexing. Samples were centrifuged at 3,000 x*g* for 4 min at 4°C, supernatant was removed and kept on ice. 300 µl clarified lysate was taken for oxidized glutathione (GSSG) analysis. 6 µl of 1:10 2-vinylpyridine (diluted in 1xKPE buffer) (Sigma-Aldrich) was added to each sample, and samples were incubated at room temperature for 1 hr. During this incubation time total glutathione analysis was carried out. After 1 hr, 18 µl triethanolamine (diluted 1:6 with 1xKPE) (Sigma-Aldrich) was added and tubes were incubated for a further

10 min at room temperature to quench the excess 2-vinylpyridine. A sample of 20 µl from each GSSG sample was pipetted in triplicate into a 96-well plate. 60 µl 5,5'-dithio-bis-(2-nitrobenzoic acid (DNTB) (0.02 M dissolved in 1xKPE buffer) (Sigma-Aldrich) and 60 µl glutathione reductase (0.333 units per ml dissolved in 1x KPE buffer) (Sigma-Aldrich) was then added to the 96-well plate. Plates were incubated at room temperature for 2 min before the addition of 60 µl β-NADPH solution (0.008 M dissolved in 1X KPE buffer) (Sigma-Aldrich). Absorbance at 412 nm was then immediately read in a SpectraMax M2 plate reader with measurements taken every 30 sec for 2 min. For the determination of total glutathione, 50 µl of sample was diluted 1:6 with 1X KPE buffer. A volume of 20 µl sample was then pipetted in triplicate into a 96-well plate and analysis carried out as before with GSSG samples. Seven-step serial dilutions from 26.4 µM to 0.4125 µM of both GSH (reduced from) (Sigma-Aldrich) and GSSG (Sigma-Aldrich) were used to generate standard curves. The concentration of total GSH and GSSG in the sample was calculated using linear regression to calculate the values obtained from the standard curve. Reduced glutathione was calculated by subtracting 2X GSSG reading from the total glutathione reading. All results were then normalised to total protein concentration in the sample.

2.6.13 Colony formation assay with crystal violet staining

To access the ability of cancer cell lines to propagate, colony formation assays were carried out. Cells were seeded at a density of 500 cells per 2 ml media/ well in 6-well plates. At 24 hr post-seeding the plates were drugged with inhibitors if required. Plates were then left for 2 weeks in a 37°C 5% CO_2 incubator. The media was then removed from the plates and the wells were washed twice with ice cold PBS. Cells were then fixed with ice cold methanol for 30 min at room temperature with gentle agitation. Methanol was then removed from the plates and the colonies stained with 5%(w/v) crystal violet (Sigma-Aldrich) for 1 hr at room temperature on a rocker. Plates were then washed repetitively in water to remove any unbound stain. Photographs of each well were taken.

2.6.14 Proliferation studies using an IncuCyte Zoom™ imagine system

The basal proliferative kinetics across different live cell lines or in the presence of inhibitors was monitored using the IncuCyte Zoom™ imaging system and software (percentage confluence) (Essen Biosciences, London, United

Kingdom). Initially, cells were seeded at a range of densities; 200, 500, 1000, and 2000 cells per well in a 96-well plate. These plates were then scanned every day for 6 consecutive days at the same time of day and the cell density that showed clear exponential growth kinetics resulting in approximately 90-95% confluence within 6 days was chosen. From the prior experiment, 1000 cells per well were seeded in 96-well plates. 24 hr after, the plates were treated with respective inhibitors as per experimental design and scanned in the IncuCyte Zoom™. The plates were maintained in a 37°C 5% CO_2 incubator in between scans and were scanned every day for 6 days. After 6 days, the percentage confluence of the wells was calculated using the IncuCyte Zoom™ software.

2.6.15 MTT

3-(4,5-Dimethylthiazol-2-Yl)-2,5-Diphenyltetrazolium Bromide (MTT) assays were carried out in a 96-well plate format to access cell metabolic activity as a measurement of cell viability in the presence or absence of inhibitors combined with platinum-based chemotherapeutic agents. Cells were seeded at 5,000 cells per 100 μl per well in triplicate in a 96-well plate. At 24 hr post-seeding the cells were treated with a six-stage cisplatin serial dilution from 25.33 μM, alone and in the presence of 5 μM MK2206, 5 μM CT99021,1 μM Harmine, 1 μM ID-8 or 100 mM BSO. After 72 hr, the media was removed and 100 μl of 1mM MTT (Sigma-Aldrich), diluted in PBS was added directly to the wells and the plates were incubated for 1 hr at 37°C at 5%CO_2. The MTT solution was removed and 100 μl DMSO was added, and the plates were further incubated in the dark for 1 hr at room temperature on a shaking platform. Absorbance's was then read using a SpectraMax M2 plate reader at a wavelength of 570 nm.

2.6.16 Subcellular fractionation – nuclear and cytosolic separation

To determine if a certain protein was located in either the nuclear or cytoplasmic compartment of the cell, subcellular fractionation was carried out. Cells were seed in 10 cm dishes to obtain a confluency of ~80-90% within 24 hr. If required cells were treated with 5 μM CT99021 for 24 hr. The media was then aspirated from the plates and the cells washed thoroughly twice in ice cold PBS, and then pelleted by centrifugation (13,000 x*g*) in 1 ml of PBS. Pelleted cells were then resuspended in 400 μl buffer A (10 mM hepes (pH 7.9), 10 mM KCL, 0.1 mM EDTA, 0.1 mM EGTA, 1 mM β-mercaptoethanol) and left on ice for 10 min. 10 μl of 10% (v/v) NP-40 was then added and the cells were lysed through

vortexting briefly. Samples were then centrifuged briefly at 13,000 x*g* for approximately 10 sec. The supernatant containing the cytoplasmic fraction and was transferred into a new eppendorf. The pellet was then washed four times by the addition of 400 µl buffer A, briefly vortexting and then centrifuging as before. After the forth wash, the pellet was resuspended in 100 µl buffer B (20 mM hepes (pH 7.9), 400 mM NaCL, 1 mM EDTA, 1 mM EGTA, 1 mM β-mercaptoethanol) and then sonicated for 10 sec. Lysates were then centrifuged at 13,000 x*g* for 15 min at 4°C. The supernatant contains the nuclear fraction. 50 mM NaF and one cOmplete™ EDTA- free Protease Inhibitor Cocktail tablet was freshly added to 10 ml aliquots of buffer A and buffer B immediately before use. A DC protein assay, and western blotting were carried out for the proteins of interest, as described above in sections 2.5.2, 2.5.3 and 2.5.5.

2.6. 17 Cyclohexamide chase

Cyclohexamide (CHX) is an inhibitor of protein biosynthesis, which acts through halting translational elongation, that is often used to determine the half-life of a protein. Cells were seeded in 6-well plates to obtain a confluency of ~80-90% after 24 hr. A volume of 10 µl of a 100 mg/ml CHX solution (diluted in media) (Sigma-Aldrich) was then added to the wells at 15 min intervals from 0 min to until 75 min. Media was then aspirated from the wells, before washing wells twice with PBS and then harvesting protein as described in section 2.5.1. A DC assay and western blotting were then carried out for the proteins of interest.

2.6.18 NQO1 activity assay

The enzymatic activity of NQO1 in cells was determined using the "Prochaska" method adapted from Fahey et al. (2004). Cells were seeded at a density of $1x10^4$ cells per 200 µl in a 96-well plate. At 24 hr post-seeding cells were drugged with a nine-step serial dilution starting at 50 nM CDDO-Me. Cells were incubated in drugs for 48 hr before analysis. Media was removed from wells by flicking contents in to the sink, plates were then washed with PBS using a BioTek microplate washer. A volume of 75 µl digitonin solution (6.5 mM digitonin, 2 mM EDTA (pH 7.8)) was added to each well and plates were then placed on a shaking platform for 20 min. An aliquot of 20 µl from each well was then transferred to a 96-well plate for protein quantification. 200 µl assay buffer (composition of which is described in the Table 2.6.2 below) was added to each well and plates were incubated for 2 min at room temperature. 50 µl of dicoumarol

solution (0.3 mM dicoumarol, 5 mM KPO_4 (pH 7.4), 0.5%DMSO) was then added before reading absorbance using a SpectraMax M2 plate reader at a wavelength of 570 nm. Readings were the normalised to total protein concentration of the sample.

Reagent	Volume (For 1x 96-well plate)
0.5 M tris/HCL, pH 7.4	1 ml
10% BSA (Sigma-Aldrich)	1 ml
1.5% Tween20	133 µl
7.5 mM flavin adenine dinucleotide (FAD) (Sigma-Aldrich)	13.3 µl
150 mM glucose-6-phosphate sodium salt (G-6-P) (Sigma-Aldrich)	133 µl
50 mM nicotinamide adenine dinucleotide phosphate (NADP) (Sigma-Aldrich)	12 µl
G-6-P-D (1,000 u/ml) (Sigma-Aldrich)	40 µl
MTT (Sigma-Aldrich)	6 mg
25 mM menadione (Sigma-Aldrich) in acetonitrile (4.3 mg/ml in acetonitrile)	20 µl

Table 2.6.2: Chemical components of the assay buffer used in the NQO1 activity assay

2.6.19 RNA extraction

For the generation of cDNA for subsequent mRNA expression analysis, RNA was extracted from cells. Cells were seeded to produce a confluency of ~80-90% on the day of harvest. Media was removed from the cells, followed by rinsing with PBS twice before adding 1-2 ml of trypsin (volume depending on the size of the plate) and incubating for 3 min at 37°C. The sides of the plates were then agitated to ensure the detachment of all cells. An equal volume of media was then added and the liquid resuspended several times before transferring to a pre-chilled eppendorf. Cells were then pelleted through centrifugation at 13,000 x*g* for 10 min in a 4°C centrifuge before lysing the pellet in 200-350 µl RLT (supplemented with 2-mercaptoethanol (2ME) (Sigma-Aldrich). The lysed pellet was then passed through a QIAShedder (QIAgen) column. RNA was then extracted according to manufacturer's guidelines using a RNAeasy Mini kit. An additional 15 min DNAse (QIAgen) digestion was carried out and the elution stage with RNA free water was repeated twice. RNA yield and purity was then measured using a using Nanodrop ND-1000 spectrophotometer.

2.6.20 Reverse transcription of RNA to cDNA

Extracted RNA was the reverse transcribed for 1 hr at 37°C to cDNA through PCR. A sample of 500 ng RNA was reverse transcribed to cDNA using

an Omniscript RT Kit (QIAgen) according to the manufacturers guidelines provided. The obtained cDNA was subsequently diluted to 50 ng/µl with RNAase free water. Samples were then stored at -20 until use.

2.6.21 Measurement of mRNA expression using real time PCR - TaqMan®

To access the variations in gene expression across different cell lines and in the presence of inhibitors, real time (RT)-PCR was carried out. RT-PCR was preformed using a QuantStudio® 5 (ThermoFisher Scientific) system. All samples were analysed in triplicate and were normalised to the expression of a housekeepering gene. All reactions were performed using 250 ng of cDNA with 0.6x TaqMan master mix (Universal Master Mix II, no UNG, (Applied Biosystems)) and in 96-well plate format (MicroAmp Optical 96-well Reaction plate with barcode, ThermoFisher Scientific) Plates were sealed using MicroAmp optical adhesive film (ThermoFisher Scientific)

The probes used were either commercially bought optimised 20X mixes (Applied Biosystems) used at a final concentration of 1x or previously optimised individual primer (900 nM) and probes (200 nM) sets. Sources and sequences of all probes can be found in Appendix 8.4.

The PCR conditions used are shown in Table 2.6.3.

Temperature (°C)	Time	Cycles
50	2 min	1
95	10 min	1
95	15 sec	40
60	1 min	

Table 2.6.3: PCR conditions for transcribing RNA into cDNA

Probes used for RT-PCR contain a fluorescent 5'FAM reporter dye and a 3' TAMRA/MGB quencher. The ratio of change in Ct value (dCt) between the control and experimental samples relative to their respective loading controls, was determined using the Pfaffl method (Pfaffl, 2001).

2.7 Bioinformatic analysis

Bioinformatics analysis of cell lines was carried out using the TIBCO Spotfire® software. Three cell lines databases: Sanger, Cancer Cell Line Encyclopaedia (CCLE) and AstraZeneca were combined in TIBCO Spotifire® and filtered for lung cancer cell lines. This provided 239 cell lines that were then further filtered for the presence of mutations in genes of interest. Cell lines

possessing a mutation in a particular gene were then further filtered for the presence of mutations in additional genes.

NRF2-target genes expression analysis was carried out by first filtering the lung cancer cell lines (239 cell lines) for the presence of mutations in a gene of interest, then the mRNA expression of NRF2-target genes was compared between the population of cell lines harbouring mutations and those that are "wildtype" for the gene in question. Gene expression values were normalized using frozen robust multiarray analysis (fRMA) (McCall et al., 2010) and subsequently scaled between 1 and 100. The log2 of this value was reported for each gene in the TIBCO Spotfire® file. Hence, the expression data ranged between 0 and 6.64. For subsequent analysis of the alterations in classical NRF2-target gene expression between; *NFE2L2*-mut, *NFE2L2*-wt (-*KEAP1*-mut), *KEAP1*-mut and *KEAP1*-wt (-*NFE2L2*-mut), only the Sanger data set was used. Average gene expression was calculated for the different groups of cell lines, and this value along with the standard error of the mean (SEM) was graphed. Bioinformatics analysis of tumour samples was preformed using cBioPortal software (Cerami et al., 2012;Gao et al., 2013). Lung cancer datasets were chosen and then filtered for the presence of mutations in certain genes. This provided the somatic frequency of mutation and the mutational mapping. The "enrichment" tab was then used to filtered for the presence of other somatic mutations. Log ratios and direction/tendency categorisation was provided by cBioPortal for the tumour data. Log ratios and direction/tendency categorization for the cell line bioinformatics data were calculated according to the cBioPortal method.

2.8 Data Analysis

Data analysis was carried out using GraphPad Prism ® 5 software. Statistical significance of the results was calculated using *t*-tests. Significant increases have been denoted throughout the text using the following symbols: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$. Significant decreases are denoted with a $ symbol, with the same cut-offs applying.

3.0 KEAP1-independent regulation of NRF2 through the $SCF^{\beta\text{-}TrCP}$/ GSK3-β axis is dependent on the presence of a "priming kinase".

3.1 Introduction

3.1.1 The Ubiquitin- proteasome system

All eukaryotic proteins are degraded in the cell through either one of two routes; extracellular proteins that enter the cell under stressed conditions are degraded by lysosomes and intracellular proteins are degraded by the ATP-dependant ubiquitin-proteasome system (UPS) under both stressed and unstressed conditions. Alterations in the UPS affect a wide range of biological processes such as cell cycle and apoptosis and are therefore implicated in a wide range of diseases, such as cancer, where impaired UPS function has been linked to poor prognosis (Shah et al., 2001).

3.1.2 Ubiquitination process

All proteins that are degraded by the proteasome first undergo labelling with the protein modifier ubiquitin in a process called ubiquitination which involves more than 1000 proteins. Ubiquitin is a small ubiquitously expressed 76-amino acid poly-peptide which is attached to the target protein though a multi-step cascade of reactions catalysed by three enzymes: ubiquitin-activating enzymes (E1s), ubiquitin-conjugating enzymes (E2s) and ubiquitin ligases (E3s) (Spataro et al., 1998). First, E1s mediate the ATP-dependent activation of ubiquitin, followed by the conjugation of ubiquitin to an intermediate protein carried out by E2s and then finally E3s carry out the ligation of ubiquitin to a Lys residue on the target protein. Ubiquitination can involve the attachment of a single molecule of ubiquitin, which is called mono-ubiquitination or attachment of a chain of ubiquitin molecules through repetitive cycles of ubiquitination; this latter process is called poly-ubiquitination. The first ubiquitin molecule is bound to a specific Lys residue in the substrate and the chain is generated through the ubiquitination of the previous ubiquitin molecule at one of its seven lysine residues (Lys-6, Lys-1, Lys-27, Lys-29, Lys-33, Lys-48 and Lys-63). Over 50% of all ubiquitin chains are formed through attachment at Lys-48 and attachment at this residue results in proteasomal degradation because poly-ubiquitin tags generated at this site can be recognized by the 26S proteasome to allow protein unfolding which is

essential for protein degradation (Robertson et al., 2018; Swatek and Komander, 2016).

The UPS is highly specific and the majority of this specificity is provided by the large family of E3 ubiquitin ligases, which can be subdivided into three main groups; Homologues to the E6AP carboxyl terminus (HECT) group, the really interesting new gene (RING) group and the RING-between RING (RBR) group. The rest of this chapter will focus on the function and regulation of the RING subgroup of E3 ligases (Robertson et al., 2018).

3.1.3 SCF family of E3 ubiquitin ligase

There are over 600 members of the RING family of E3 ubiquitin ligases that all catalyse the direct transfer of ubiquitin from the E2 enzyme to the substrate. Members of the RING subfamily can function as monomers, dimers or large multi-subunit complexes, such as the Skp, Cullin, F-box containing (SCF)-complex of which there are 69 found in mammals (Gorelik et al., 2016). The SCF complex is composed of four components: an S-phase-kinase-associated protein-1 (SKP1), cullin-1 (CUL1), RBX1 and an interchangeable F-box protein. The F-box protein is the most variable component in the whole SCF complex and confers a high degree of specificity to the complex. Each of the four proteins are involved in different binding interactions to form the complex. RBX1 functions to bind to both CUL1 and the E2-conjugating enzyme, while CUL1 functions as a scaffold to join RBX1 to SKP1. SKP1 is composed of two domains, the N-terminal of which binds to CUL1 and the C-terminal domain that binds to the N-terminal of the F-box protein through an approximately 40-amino acid long F-box motif. The C-terminal region of the F-box protein binds substrates through protein: protein-interaction domains such as WD40 repeats that adopt a β-propeller structure, which binds specifically to phosphorylated substrates and leucine-rich repeats (LRR) that have a α/β repeats and bind substrates independently of phosphorylation status (Robertson et al., 2018; Wu et al., 2003; Metzger et al., 2012; Buetow and Huang, 2016; Pickart, 2001).

3.1.4 $SCF^{\beta TRCP}$ complex

Beta-Transducin repeat containing protein (β-TrCP) is a classical highly conserved example of an F-box protein, with seven C-terminal WD40 repeats that bind target substrates. There are two mammalian isoforms of β-TrCP, called

β-TrCP1 and β-TrCP2, which appear to be functionally redundant and therefore will be referred to in this thesis as only β-TrCP.

β-TrCP functions as an E3 ligase to a wide range of substrates which are involved in a variety of cell signalling pathways including; β-catenin, IKβ, SMAD3 and Wee1, and hence dysregulation in β-TrCP is implicated in several diseases, such as cancer. However, one aspect that all β-TrCP substrates have in common is they all contain a destruction motif to which β-TrCP binds based around the sequence, DSGØXS, where both serine's are phosphorylated, Ø represents a hydrophobic residue and X is any amino acid. When in its phosphorylated state, this motif is often referred to as the phosphodegron (Wu et al., 2003; Seo et al., 2009).

3.1.5 Interplay between GSK-3 and β-TrCP

Due to the spacing between the two phosphorylated Ser residues in the DSGØXS β-TrCP destruction motif being similar to the GSK-3 substrate consensus sequence, Ser/Thr XXX(X) Phospho (Ser/Thr), it is to be expected that β-TrCP substrates are often first phosphorylated by GSK-3 (Figure 3.1.1). One such example is the classic β-TrCP substrate β-catenin. In this case, only when GSK-3 phosphorylates Ser-33, Ser-37 and Thr-41 of β-catenin is β-TrCP able to bind β-catenin resulting in its subsequent degradation. This suggests that GSK-3 phosphorylation regulates β-TrCP recognition and hence protein stability. Other examples of proteins phosphorylated by GSK-3 before subsequent β-TrCP mediated proteasomal degradation can occur include; Snail, MYC and many more (Liu et al., 2002; Gao et al., 2014 ;Robertson et al., 2018).

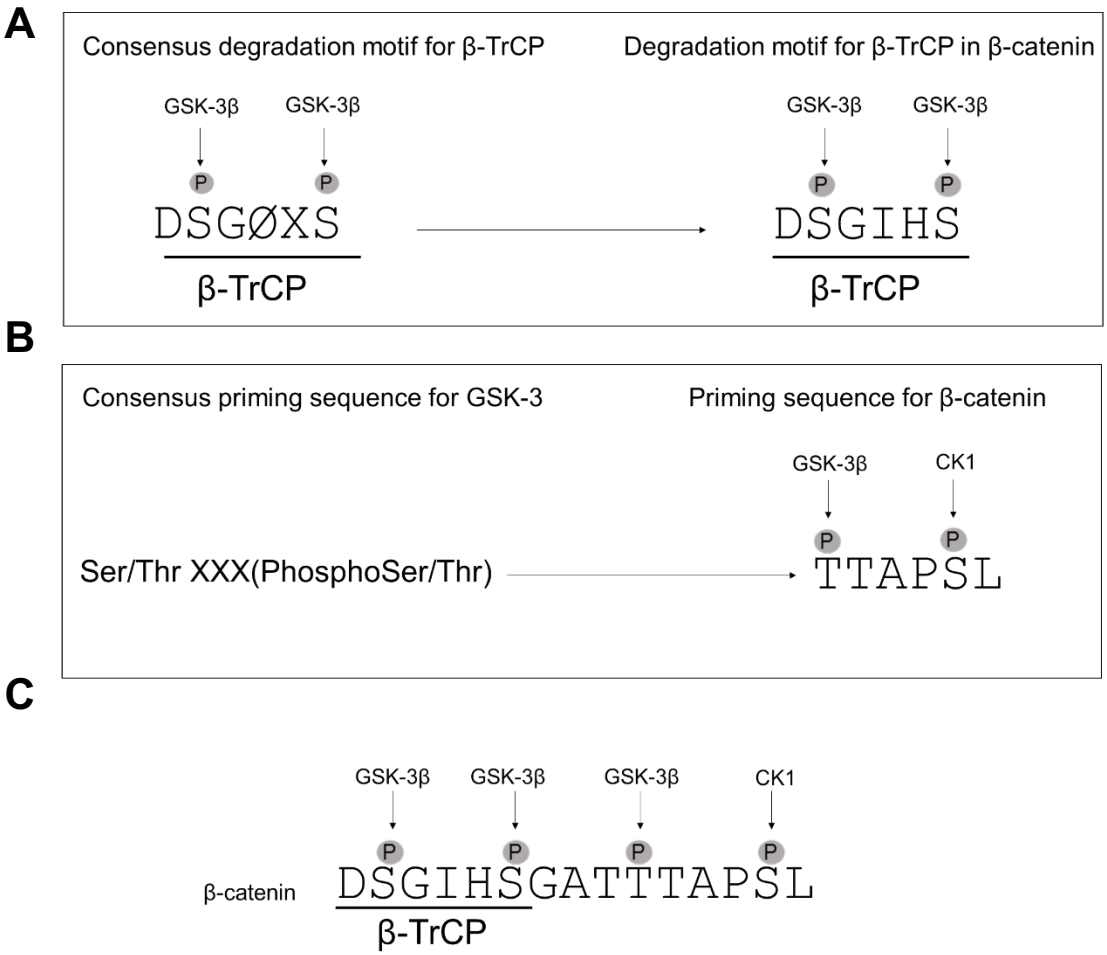


Figure 3.1.1:The destruction sequence for β-TrCP resembles the consensus priming sequence for GSK-3

(A)The degradation motif for β-TrCP in β-catenin reflects the prototypic consensuses degradation motif. The bold line under the sequence indicates where β-TrCP binds and the amino acid sequence displayed corresponds to human β-catenin. Phosphorylated residues are indicated with grey circles above them. The kinase phosphorylating at that site is indicated with the arrows above. (B) The priming sequence found in β-catenin conforms to the classical consensus priming sequence for GSK-3. Phosphorylated residues are indicated with grey circles above them. The kinase phosphorylating at that site is indicated with the arrows above. The sequence displayed corresponds to human β-catenin. (C) CK1 functions as a priming kinase phosphorylating β-catenin five amino acids C-terminal from the first GSK-3β phosphorylation site. This priming event enhances subsequent GSK-3β phosphorylation at three separate amino acids. Phosphorylation by GSK-3β forms the phosphodegron to which β-TrCP is recruited. The bold line under the sequence indicates where β-TrCP binds and the amino acid sequence displayed corresponds to human β-catenin. Phosphorylated residues are indicated with grey circles above them. The kinase phosphorylating at that site is indicated with the arrows above.

3.1.6 GSK-3 signalling

GSK-3 is a Ser/Thr kinase, so named after its ability to phosphorylate the substrate glycogen synthase, is ubiquitously expressed in human tissues, with some degree of tissue to tissue variability. Since the discovery of the involvement of GSK-3 in glycogen metabolism in 1980 (Cohen and Frame, 2001; Embi et al., 1980) it has been found to be involved in a wide range of other cellular processes targeting over 100 proteins in the cell (including the transcription factors AP-1 and CREB), and due to this diverse range of substrates dysregulation of GSK-3 signalling has been implicated in a wide range of diseases, such as, but not limited to: bi-polar mood disorder, Alzheimer's disease, diabetes and cancer (Jope and Johnson, 2004). There are two isoforms of GSK-3 present in humans, GSK-3α and GSK-3β, which appear to have very similar characteristics, apart from GSK-3α having a slightly heavier molecular mass of 51 kDa, in comparison to GSK-3β which runs at 47 kDa on a SDS-PAGE (Sutherland, 2011).

GSK-3 is unusual in comparison to other kinases in two aspects; firstly, that it is inhibited under growth factor stimulation and secondly that it requires its substrates to have already undergone a phosphorylation event by another protein kinase.

3.1.7 Regulation by GSK-3 signalling

Several different protein kinases have been shown to regulate GSK-3 activity through phosphorylation of the N-terminal Ser-21 in GSK-3α and Ser-9 in GSK-3β. GSK-3 is also found to be associated with multiprotein complexes, such as the axin-APC complex, that when altered regulates GSK-3 activity. Also, GSK-3 has been implicated to regulate its own activity through autophosphorylation of Tyr-279 and Tyr-216, in GSK-3α and GSK-3β respectively (Cole et al., 2004).

One of the major breakthrough findings in terms of GSK-3 regulation was found by Cross and colleagues who demonstrated that AKT (also known as protein kinase B; PKB) directly phosphorylates GSK-3α at Ser-21 and GSK-3β at Ser-9, and that both these phosphorylation events results in inhibition of GSK-3 activity (Cross et al., 1995). This inhibitory phosphorylation leads to a blocking of the substrate binding site. This is of particular interest as the AKT signalling pathway is extremely important for cell survival and is one of the most commonly deregulated pathways in cancer.

Amino acid stimulation that leads to activation of ribosomal protein S6 kinase (S6K), growth factor stimulation leading to activation of the Ras/Raf pathway and active PKC signalling also lead to the inhibitory phosphorylation of GSK-3 through the same sites that are phosphorylated by AKT (Cohen and Frame, 2001).

3.1.7.1 Autophosphorylation at Tyr279/Tyr216 in GSK3α/GSK-3β

Tyr-279 in GSK-3α and Tyr-216 in GSK-3β are both located in the activation loop and have been shown to be phosphorylated under resting conditions. It is not unlikely that the phosphorylation at these sites impacts GSK-3 activity as these residues are located in a very highly conserved region and phosphorylation at the equivalent position in mitogen-activated protein kinase (MAPK) has been shown to dramatically enhance MAPK activity (Hughes et al., 1993). However, the function of this phosphorylation event and also the kinase that mediates it, is still under much debate. Work by Dajani and colleagues on the crystal structure of GSK-3β, suggested that Tyr-216 phosphorylation is not that influential in terms of regulating GSK-3β activity. However, they do show that when Try-216 is unphosphorylated it partially blocks the binding pocket to which the phosphate on the primed substrate (see below for explanation of priming) binds (Dajani et al., 2001). Therefore, having Tyr-216 phosphorylated may enhance binding but is not a requirement for binding and may have more of a pronounced effect in primed substrates. They do not elude to the protein kinase responsible. In contrast to this work, Cole and colleagues have reported that phosphorylation at both Tyr-279 in GSK-3α and Try-216 in GSK-3β is an autophosphorylation event mediated by GSK-3 itself, but they did not shown any data to indicate the function of this phosphorylation (Cole et al., 2004).

3.1.7.2 Axin-APC complex

WNT signalling is fundamental to cell growth and differentiation. In resting cells GSK-3β phosphorylates (in concert with Casein kinase 1; CK1) several serine residues in β-catenin which is fundamental for the recruitment of $SCF^{\beta\text{-}TrCP}$ complex, which leads to the ubiquitination and destruction of β-catenin by the proteasome. Therefore, GSK-3β functions as a negative regulator of WNT signalling. When the ligand WNT is present GSK-3β is inhibited meaning β-catenin can perturb proteasomal degradation. The phosphorylation of β-catenin by GSK-3β occurs in a multiprotein complex composed of: axin, adenomatous

polyposis coli (APC), β-catenin, CK1 and GSK-3β (Liu et al., 2002; Pronobis et al., 2015). The complex potentially functions to bring GSK-3β physically closer to β-catenin but, also interestingly when in this complex GSK-3 appears to be somewhat protected from other forms of GSK-3 regulation, such as inhibitory phosphorylation by the protein kinase AKT (Doble and Woodgett, 2003).

3.1.8 GSK-3 specificity- the requirement for a priming kinase

The phenomenon of pre-phosphorylation of a substrate by another kinase before GSK-3 can phosphorylate is called “priming” and is carried out by a “priming kinase”. The best characterised GSK-3 substrates have been shown to require priming by several different protein kinases. Some examples of primed GSK-3 substrates include; glycogen synthase which is primed by casein kinase 2 (CK2) β-catenin which is primed by CK1 and eukaryotic initiation factor 2B (eIF2B) which is primed by DYRK2. It is important to identify and study these priming kinases as the presence of priming phosphorylation event enhances GSK-3 affinity towards the primed substrate by over 90% (Sutherland, 2011; Robertson et al., 2018). All priming phosphorylation events take place at an amino acid either 4 or 5 residues C-terminal from the proposed GSK-3 phosphorylation site, conforming to the consensus sequence Ser/Thr XXX(PhosphoSer/Thr). Crystal structure studies have releveled that GSK-3 has a phosphate binding pocket that binds to the primed substrate positioning it closer to the GSK-3 catalytic site (Dajani et al., 2001). It should be mentioned that there are some GSK-3 substrates that do not require priming but the mechanism by which GSK-3 targets these substrates is still poorly understood.

3.1.9 The role of GSK-3 in disease

GSK-3 is at the core of several important biological pathways and hence alterations in its signalling has been implicated in a wide range of diseases, of which three are discussed below.

3.1.9.1 GSK-3 and Alzheimer’s disease

Alzheimer’s disease (AD) is a chronic and progressive form of dementia with an exponentially increasing prevalence worldwide (Qiu et al., 2009). The symptoms of AD are extremely difficult to diagnose in the early stages of the disease and are subsequently often overlooked. The late stage of the disease presents with extremely debilitating cognitive defects. The two main forms of pathophysiology associated with the disease are the formation of cleaved β-

amyloid peptides and aggregates of the microtubule associated protein, tau (Giese Karl, 2009). A small percentage of AD cases have a familial link, and in these instances, patients often possess mutations that lead to altered WNT signalling and a subsequent increase in GSK-3 activity. GSK-3 has also been linked to the development of sporadic AD, as GSK-3α is involved in the degradation of β-amyloid peptide but also the accumulation of this peptide has been shown to influence the activity of GSK-3β; this may in part enhance the GSK-3β mediated hyperphosphorylation of tau (Hernández et al., 2009). Over 30 phosphorylation sites on tau have been identified to be potentially phosphorylated by GSK-3β. Hyperphosphorylated tau forms aggregates in axons and dendrites leading to diminished synaptic transmission and ultimately neurodegeneration.

The generation of isoform specific and neuronal-specific GSK3 knock out mice has greatly aided the understanding of the role this kinase plays in the development of neurological disorders. Since post-mortem brains of AD patients show elevated activity of GSK-3β the majority of the mouse models of AD have focused on this isoform. Neuronal specific GSK-3β knock out mice show impaired function in memory tasks and also smaller dentate gyrus, which is the area of the brain responsible for memory (Gomez-Sintes et al., 2011; Kondratiuk et al., 2013). Also knock down of both isoforms of GSK-3 leads to reduced tau phosphorylation resulting in less aggregate formation, in transgenic AD models (Hurtado et al., 2012).

There is still much more to learn about the role GSK-3 plays AD but research is limited due to the lack of healthy and diseased human tissue (Kremer et al., 2011).

3.9.2 GSK-3 and diabetes mellitus

GSK-3 has often been linked with diabetes since the initial discovery of the ability of this kinase to regulate insulin signalling through phosphorylation of glycogen synthase, the key enzyme involved in glycogen metabolism. The inhibition of glycogen synthase in muscle tissue has been linked to the development of DM, with type 2 diabetic human patients showing a 2-fold induction of GSK-3 activity (Clodfelder-Miller et al., 2005). DM is a disease characterised by the impaired production and utilization of glucose, which is currently increasing in prevalence in the western world. Murine models of Type 2

DM have shown enhanced GSK-3β activity in the tissues of high-fat diet fed mice that are insulin insensitive. Also, GSK-3 has been implicated to be involved in the translocation of the principle glucose transporter, GLUT4, from the intracellular space to the plasma membrane (Chong et al., 2012). GSK-3 has been implicated to not only be hyperactivated in peripheral tissues leading to diabetes pathophysiology but also has been implicated to be associated with the cognitive impairment that is found in several diabetes patients. Studies have shown that insulin signalling in the brain leads to activation of the AKT pathways which results in the inhibition of GSK-3 activity and impaired glucose homeostasis (Jope et al., 2007). This supports the idea that there is several shared disease traits between diabetes and AD, with AD often being referred to as type 3 diabetes (Chong et al., 2012).

3.1.9.3 GSK-3 and cancer

Several proteins that play a role in key cancer pathways are required to be phosphorylated by GSK-3 for subsequent degradation through the $SFC^{\beta\text{-}TrCP}$ complex, but even though GSK-3 does play a role in cancer it is still under much debate whether that role is tumour promoting or tumour suppressive. GSK-3 overexpression is seen in several cancer tumour types e.g. colon and ovarian and is thought to provide a tumour suppressor role by inhibiting the WNT signalling pathway through phosphorylation of β-catenin. Also, several cancer associated mutations have been shown to occur in GSK-3 substrates at the GSK-3 phosphorylation or priming sites, this reduces GSK-3 mediated phosphorylation and hence the mutated protein evades $SCF^{\beta\text{-}TrCP}$-mediated proteasomal degradation and accumulates in the cell. For example, mutation in the GSK-3 phosphorylation sites on β-catenin and c-Myc leads to an accumulation in these oncogenic proteins. GSK-3 is also a key negative regulator of the WNT, Notch and Hedgehog pathways, which are all heavily implicated in the development of cancer (McCubrey et al., 2014). The development and progression of GSK-3 inhibitors for treatment of other diseases, such as AD and DM, is often halted due to concerns that their administration may lead to the development of cancer (Patel and Woodgett, 2008). Although this may be true, the data supporting this argument are extremely variable. However, recently published data highlights the potential for GSK-3 inhibitors to actually be beneficial in cancer treatment (Wang et al., 2008b; Ugolkov et al., 2016).

3.1.10 Development of reagents to study GSK-3: mouse models and inhibitors

To gain a greater understanding of the complex role that GSK-3 plays in cell biology, isoform-specific mouse models and a range of targeted inhibitors have been developed.

3.1.10.1 Mouse models

A key breakthrough in understanding the differences between the two GSK-3 isoforms were the generation of isoform specific GSK-3α and GSK-3β knock out mice. Genetic knockout of the alpha isoform produced viable animals with no adverse phenotype. GSK-3α mice surprisingly showed increased glucose tolerance and insulin sensitivity, in comparison to wildtype matched control animals (MacAulay et al., 2007). MacAulay and colleagues attribute this to a reduction in plasma leptin levels. Interestingly the phosphorylation of glycogen synthase was unaltered in these animals, suggesting potentially different isoforms of GSK-3 mediate the phosphorylation of different target substrates. In contrast to the viable GSK-α specific knockout mice, GSK-3β knockout mice die in gestation due to severe liver damage (Hoeflich et al., 2000). Hoeflich and colleagues attribute this fatality to increased NFκB signalling leading to increased TNFα production. To understand more about the role that GSK-3β plays in a viable mouse model, tissue-specific GSK-3β knockout mice were generated (Patel et al., 2008). Genetic ablation of GSK-3β in the liver produced viable mice with no notable phenotype. However, skeletal muscle specific GSK-3β knockout mice were also viable but display enhanced glucose tolerance. This datum suggests that there may be tissue specific effects of GSK-3 isoforms.

3.1.10.2 GSK-3 inhibitors

GSK-3 is at the heart of several key pathways that are involved in several diseases, making it a useful therapeutic target. Lithium is a potent well characterised inhibitor of both GSK-3α and GSK-3β isoforms that has been used as the first line treatment of mood related disorders, such as bipolar disease, worldwide for decades (Freland and Beaulieu, 2012). Even though lithium is fairly specific it has been shown to inhibit other kinases (such as the CDK which shares sequence identity with GSK-), other non-kinase targets and also needs to be used at a very high concentration to completely inhibit GSK-3 activity (Bain et al.,

2007). There are currently over 60 GSK-3 inhibitors available for the treatment of a wide range of diseases, from cancer to muscle wasting disorders (Robertson et al., 2018). The majority of compounds, like most protein kinase inhibitors, are ATP competitive meaning that they will inhibit all GSK-3 substrates at once and this could be detrimental to cell health. The fact that GSK-3 is involved in so many key biological processes raises the issue of selectivity, for example inhibition of GSK-3 may be beneficial in one disease e.g. treatment of Alzheimer's disease setting but detrimental in another e.g. treatment of cancer. This in combination with the embryonic lethality of the GSK-3β knockout mice have halted the development of new more selective GSK-3 inhibitors. However, due to the unique requirement of GSK-3 for primed substrates there is the potential to develop inhibitors that target the binding site for the primed phosphate. So far, no drugs have been developed to successfully target this site, but phosphopeptides binding to this region show great potential (Cohen and Goedert, 2004).

3.1.11 KEAP1-independent regulation of NRF2 protein stability and localization through the Neh6 domain

Aberrant expression of NRF2 is found in several cancer types such as: liver, head and neck, ovary and lung, and its over-expression associated with poor prognosis (Shibata et al., 2008). In this case, it is thought that elevation of NRF2 signalling provides a chemo-resistant and highly proliferative ability to the cancer cell. There are several routes by which expression of the transcription factor can be increased; these include epigenetic silencing of *KEAP1* through hypermethylation of the promotor region (Reviewed in (Guo et al., 2015)), accumulation of other ETGE-containing KEAP1-binding proteins indirectly influencing NRF2 through competing for KEAP1 (Hast et al., 2013), cysteine modification by oncometabolities (e.g. fumarate targeting Cys-151 and Cys-288 of KEAP1) (Ooi et al., 2011) and somatic mutations in *KEAP1* or *NFE2L2* genes.

Over 20% of lung cancer tumours have constitutive activation of the transcription factor due to mutations in either the negative regulatory protein, *KEAP1* or *NFE2L2* (Shibata et al., 2008). Somatic mutations throughout *KEAP1*, leading to a loss of function and "hot-spot" mutations in the DLG and ETGE motifs in *NFE2L2*, allow NRF2 to evade KEAP1-directed ubiquitination and degradation, leading to increased expression of cytoprotective genes. Tumours harbouring *KEAP1* and *NFE2L2* mutations are highly proliferative and intrinsically resistant

to chemotherapeutics. Therefore, understanding KEAP1-independent mechanisms by which NRF2 is degraded provides an alternative way to decrease NRF2 levels when *KEAP1* is mutated.

It has been suggested several times that there must be a separate form of NRF2 regulation other than the canonical KEAP1 degradation pathway. Work by McMahon et al (2004), found that when the Neh2 domain of NRF2 is removed the full-length mouse protein can still be degraded and only when both the Neh2 and Neh6 domain are both removed does the protein significantly stabilise. They highlighted that this redox-independent Neh6 domain, corresponding to amino acids 329 to 339 in mouse NRF2, functions completely autonomously of KEAP1 and may be critical to the regulation of NRF2 under extreme conditions of oxidative stress. Under these conditions KEAP1, which is predominantly cytoplasmic, would be rendered non-functional due to modification of its Cys residues, leading to a conformational change in the protein. Such modifications of KEAP1 results in the stabilization and translocation of NRF2 to the nucleus. McMahon et al., also postulate that potentially regulation through the Neh2 domain and Neh6 domain can occur simultaneously but in different cellular compartments (McMahon et al., 2004).

The link between NRF2 and GSK-3β was forged by Salazar and colleagues who through studying the subcellular localization of the transcription factor demonstrated that NRF2-acitivty is inhibited through phosphorylation by GSK-3β. GSK-3β was shown to negatively regulate NRF2-activity though stimulating its nuclear export (Salazar-Roa et al., 2006). This supports the previous finding that the NRF2 orthologue SKN-1 is inhibited by GSK-3β (An et al., 2005). At this point the exact GSK-3β phosphorylation site(s) was unknown but the potential role of a priming kinase was suggested.

Subsequent papers reinforced this linked between NRF2 and GSK-3β through the study of neurological diseases. Short term oxidative stress, through exposure to H_2O_2, was documented to lead to nuclear accumulation of NRF2 but long term H_2O_2 induced oxidative stress was shown to enhance cytosolic NRF2 levels through stimulation of its nuclear export (Rojo Ana et al., 2008; Rojo et al., 2008).

A couple of years after McMahon and colleagues were making their observations a paper was published by Rada et al., supporting these findings and expanding on NRF2 regulation through the Neh6 domain. Rada et al., elegantly linked that fact that several transcription factors, such as Snail and β-Catenin are regulated by GSK-3 and the presence of an evolutionarily conserved cluster of Ser residues in the Neh6 domain to suggest that GSK-3 may phosphorylate the Neh6 domain. GSK-3 depletion through either chemical inhibition or siRNA gene knockdown was shown to enhance NRF2 protein levels, independently of KEAP1. Rada et al., went on to identify a putative β-TrCP binding motif, DSGIS, within the Neh6 region which GSK-3 phosphorylates. GSK-3 phosphorylation of Ser-335 and Ser-338 was highlighted to lead to the formation of a phosphodegron motif which is subsequently recognized by β-TrCP. They also supported the idea from the McMahon paper that the magnitude of oxidative stress may influence the mechanism by which NRF2 is degraded in the cell (Rada et al., 2011). A subsequent paper by the same group (Rada et al., 2012) carried out several Nuclear Magnetic Resonance (NMR) and docking experiments to further validate that the Neh6 domain can be bound by β-TrCP specifically through Ser-335 and Ser-338. Then through the generation of several site mutants against the clusters of serine residues in the Neh6 domain, they determined the following; Ser-351 and Ser-355 are not phosphorylated by GSK-3β, GSK-3 phosphorylates Ser-335 and Ser-338 greater in the presence of intact Ser-342 and Ser-347 and phosphorylation of Ser-342 and Ser-347 still occurs in the absence of GSK-3 suggesting other kinases may also phosphorylate the Neh6 domain. Rada and colleagues hypothesized that Ser-342 and Ser-347 could potentially function as a priming site(s) for GSK-3 activity. However, at this point they could not provide any data or suggest any kinases that may phosphorylate these sites (Rada et al., 2011).

The interaction between β-TrCP and the Neh6 domain was further substantiated by Chowdhry and colleagues who identified two non-identical regions within the Neh6 located at mouse amino acids 329-339 and 263-379, that are bound by β-TrCP (Chowdhry et al., 2013). Both these regions function to control the turnover of NRF2 completely independently of KEAP1. One motif has the sequence DSGIS and the other DSAPGS, with the former being regulated by GSK-3 and the latter not. This enforces the idea that GSK-3 is central to KEAP1-

independent degradation of NRF2 by phosphorylating several key Ser residues in the Neh6 domain leading to the formation of the phosphodegron to which β-TrCP binds.

Whilst it is clear that GSK-3 and β-TrCP are capable of exerting KEAP1-independent repression of NRF2, the notion that a priming kinase is necessary in order that GSK-3 can create a phosphodegron in the Neh6 domain of NRF2 to which β-TrCP binds has not been researched. The data in this chapter identifies the specific priming site(s) located in Neh6, gives examples of kinases that may prime NRF2 for phosphorylation by GSK-3, and demonstrates the impact of priming on NRF2 degradation in the cell.

3.2 Results

3.2.1 The Neh6 domain of NRF2 contains both β-TrCP binding sites and two potential priming sites and is highly conserved across several species

NRF2 is principally controlled at a protein stability level through 26S proteasomal degradation. The Neh2 domain, spanning amino acids 1-87 in mouse Nrf2, contains both motifs to which KEAP1 binds. Under homeostatic conditions KEAP1 binds NRF2, recruits a CUL3 ligase, which ubiquitinates NRF2 and labels the transcription factor for subsequent degradation by the 26S proteasome. The Neh6 domain of NRF2, spanning amino acids 328-380 of mNRF2, contains two motifs to which β-TrCP binds. These motifs are DSGIS and DSAPGS, which are highly conserved across several species (Chowdhry et al., 2013 ;McMahon et al., 2004). The DSGIS motif has been shown to require phosphorylation at Ser-335 and Ser-338 (numbering referring to mouse NRF2 (mNRF2)) to ensure maximal binding by β-TrCP (Rada et al., 2012; Rada et al., 2011). By contrast, the binding of β-TrCP to the DSAPGS motif does not seem to be influenced by phosphorylation (Chowdhry et al., 2013) (Figure 3.2.1).

The DSGIS motif is flanked on its C-terminal side by Ser-342 and Ser-347, both of which are located immediately adjacent to a Pro residue (i.e. Pro-343 and Pro-348) within a region that is highly conserved across species. The phosphorylation status of these two amino acids has been reported to influence the ability of GSK-3 to phosphorylate Ser-335 and Ser-338 (Rada et al., 2012; Rada et al., 2011), suggesting that phosphorylation of the DSGIS motif occurs in a progressive fashion. It is well known that GSK-3 has a specificity for substrates that have undergone a prior phosphorylation event mediated by another kinase

at an amino acid positioned either 4 or 5 amino acids C-terminal from the initial GSK-3 phosphorylation site. As shown in Figure 3.2.2, we propose that phosphorylation of NRF2 at Ser-342 and Ser-347 is mediated by a currently unknown kinase and functions as a priming event for subsequent phosphorylation by GSK-3β at Ser-335 and Ser-338.

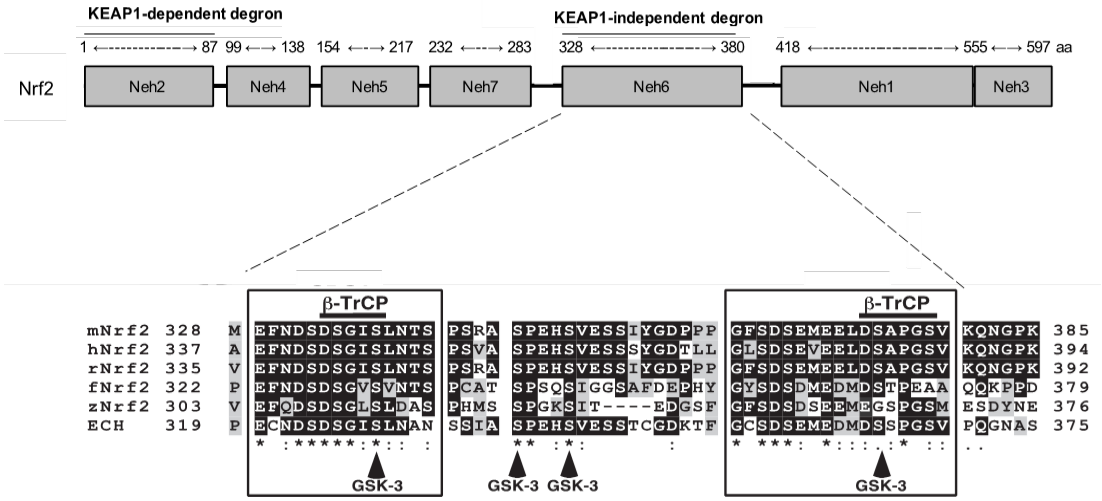


Figure 3.2.1: The Neh6 domain of NRF2 contains both β-TrCP binding sites and is highly conserved across several species.

Schematic representation of the protein structure of mNRF2. Both KEAP1-dependent and KEAP-independent degradation regions are highlighted above in bold font, also the size of each Neh domain of mNRF2 is provided by the numbers between the dotted arrows. Underneath in the expanded Neh6 domain are the amino acid sequences corresponding to the Neh6 domain of NRF2 from mouse (m), human (h), rat (r), frog (f), and zebrafish (z), along with that of ECH (that is, chicken NRF2) that have been aligned using the T-Coffee tool (at http://www.ebi.ac.uk/Tools/msa/tcoffee/). White letters on a black background represent residues that are identical across at least half of the species studied, and black letters on a grey background are conserved residues. The two black outlined boxes contain the sequences enriched with Ser and Asp residues and are the regions to which β-TrCP binds. Within the black boxes, a solid horizontal bar is shown above sequences that represent putative β-TrCP-binding sites. The residues that are predicted to be phosphorylated by GSK-3, based on the Scansite program (at http://scansite.mit.edu), are indicated at the bottom as vertical arrows.

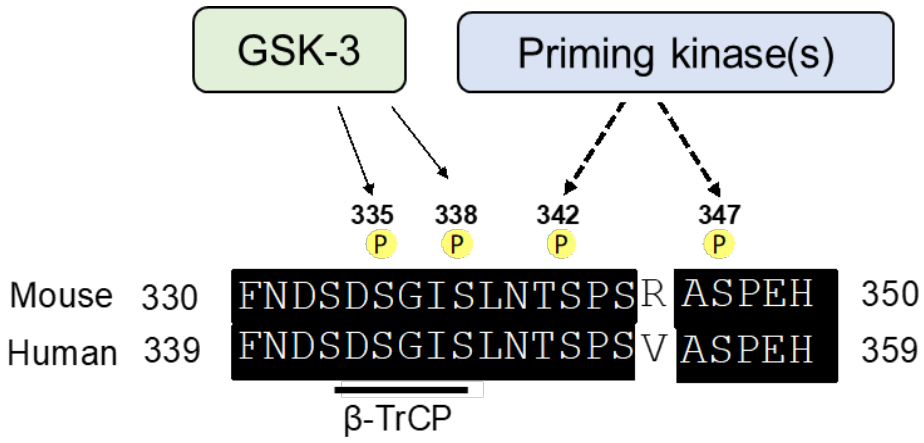


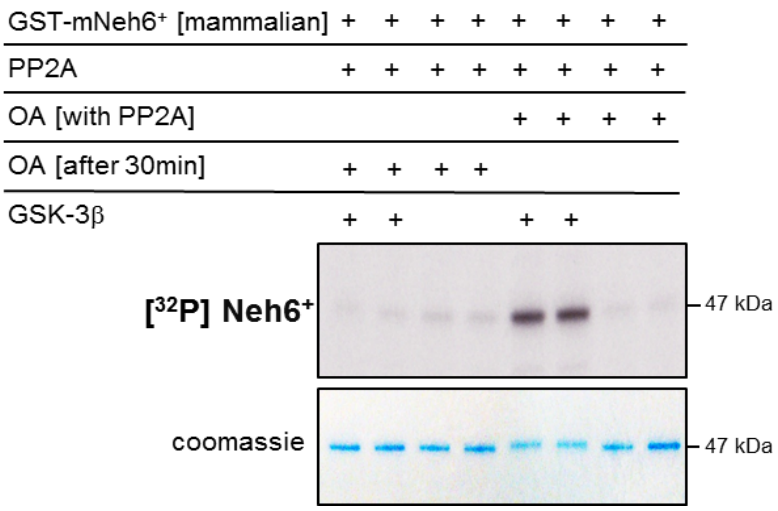
Figure 3.2.2: Potentially there are two distinct phosphorylation sites, Ser-342 and Ser-347, within the Neh6 domain which may function as priming sites for subsequent phosphorylation of Ser-338 and Ser-335 by GSK-3.

Sequence alignment of the region within the Neh6 domain of mouse and human NRF2 that contains the site to which β-TrCP binds. It also shows the GSK-3 phosphorylation site(s) and the potential priming site(s). White letters on a black background represent residues that are identical between the two species. Black letters on a white background, highlight residues which are not conserved across the two species. The horizontal bar under the residues corresponds to the β-TrCP binding region. The residues that are predicated to be putative priming kinase sites, Ser-342 and Ser-347 (numbering corresponding to the mouse NRF2), based on the Scansite program (at http://scansite.mit.edu) are indicated.

3.2.2 GSK-3β activity towards the Neh6 domain of mNRF2 is enhanced in the presence of an additional kinase

To determine whether GSK-3β mediated phosphorylation of the Neh6 domain is enhanced by the presence of a prior phosphorylation event, we exploited the fact that mammalian cells express several protein kinases. Therefore, a Neh6-containing protein expressed in mammalian cells will have been exposed to several kinases, some of which might modify the protein. A construct corresponding to amino acids 290-410 of mNRF2 (a region that will now be referred to as the mNeh6^{+} region) with a C-terminal GST tag (referred to a GST-mNeh6^{+}) was generated by Dr. Sudhir Chowdry. This construct was then expressed and purified from monkey kidney fibroblast-like COS1 cells. We then treated the purified GST-mNeh6^{+} with protein phosphatase 2A (PP2A, a Ser/Thr specific phosphatase), either with or without Okadaic acid (OA, which is an inhibitor of PP2A) before adding GSK-3β and [^{32}P] ATP to the incubation mixture. As shown in Figure 3.2.3, the incorporation of [^{32}P] phosphate into GST-mNeh6^{+} by GSK-3β was greatly accentuated in the presence of both PP2A and OA, relative to PP2A alone, suggesting that prior phosphorylation of the Neh6^{+} region by a kinase in COS1 cells enhances subsequent GSK-3β phosphorylation. Although, it should be noted that there are eight serine residues and five threonine residues within the GST tag sequence that could be potentially phosphorylated by GSK-3. Therefore, a purified GST tag should be included as a control. This experiment was carried out by Dr. Sudhir Chowdry.

A



B

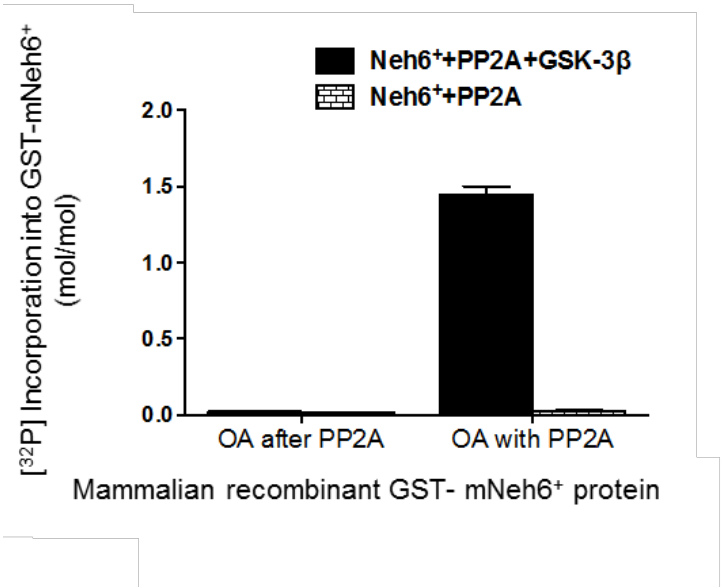


Figure 3.2.3: GSK-3 activity towards mNeh6⁺ is diminished in the presence of a phosphatase suggesting that additional phosphorylation events effect GSK-3 phosphorylation of mNRF2.

(A) To determine if the activity of GSK-3 towards the mNeh6^{+} region can be enhanced by other phosphorylation events, an in vitro kinase assay was set up with/without okadaic acid and in the presence or absence of PP2A. Two sets of reactions were setup. In the first set, the purified recombinant GST-tagged mNeh6^{+} protein (0.1 µg/µl, substrate), generated from mammalian (COS1) cells, was incubated with PP2A (2 mU) at 30°C for 30 minutes. Following the incubation period, OA (1 µM) was added and the reaction mix was incubated at 30°C for further 10 min, to inhibit PP2A. In the second set of reactions, the purified recombinant GST-tagged mNeh6^{+} protein (0.1 µg/µl, substrate) was incubated with PP2A (2 mU) and OA (1 µM) together at 30°C for 30 minutes. Following this, purified activated recombinant GSK-3β (6m U/µl (kinase)) was added to both the reaction mixtures in the presence of 10 mM $MgCl_2$ and 0.1 mM [γ ^{32}P]-ATP, and this was then further incubated for 120 min at 30°C. The reactions were terminated by addition of SDS gel loading buffer, the samples were loaded on a SDS-PAGE gel, and the excess [γ-^{32}P]-ATP was removed by electrophoresis. Gels were then stained with coomassie stain, dried and subjected to autoradiography. After which, protein bands were excised and the incorporated radioactivity was determined by scintillation counting. This image is one representative of three independent biological replicates. (B) Scintillation counts from (A). Data displayed are mean value from three independent biological replicates and the associated standards error of the mean (SEM). Experiment carried out by Dr S. Chowdhry.

3.2.3 Several members of the CMGC family protein kinases have the potential to phosphorylate the Neh6 domain of NRF2

We identified protein kinases that may prime the Neh6 region of NRF2 by conducting an *in vitro* kinase screen, in collaboration with Dr James Hastie at DSTT, University of Dundee. In these experiments we examined the relative ability of kinases to phosphorylate a region of NRF2 that included the DSGIS motif, residues that GSK-3β can phosphorylate and where a potential priming kinase may target (mNRF2 amino acids 330-350 and hNRF2 amino acids 339-359) from amongst a panel of 140 protein kinases; the enzyme panel examined had previously been used to identify small molecule inhibitors (Hastie et al., 2006; Bain et al., 2007). The capability of each kinase to phosphorylate the mouse and hNRF2-based 21-mer peptides was compared relative to its ability to phosphorylate a well-characterised standard peptide or protein (data not shown). From amongst this panel of kinases, Table 3.2.1 shows those that had the highest relative phosphorylation towards the NRF2-based peptides to background. Several kinases identified in the top 20 are members of the CMGC family of protein kinases. This is not to be unexpected as both Ser-342 and Ser-347 are flanked C-terminally by a proline residue, which is a requirement for CMGC kinases. Fifty-seven human CMGC kinases exist, which are highly evolutionarily conserved and have previously been shown to function as priming kinases for GSK-3 (e.g. CK2 primes β-catenin and DYRK primes eIF2B (Varjosalo et al., 2013; Robertson et al., 2018)). Interestingly, in our screen DYRK2 and DYRK3 were ranked first and sixth respectively in terms of their ability to phosphorylate the two 21-mer peptides, showing a 5-fold higher phosphorylation of the mouse peptide in comparison to the human peptide. Besides DYRK2 and DYRK3, TTK, MLK, CDK9 and TBK1 also extensively phosphorylated the two 21-mer peptides. Because two of the DYRK family members exhibited high activity towards the 21-mer peptides, along with the known ability of DYRK isoenzymes to prime substrates for GSK-3β, we decided to investigate further whether these enzymes can prime the Neh6 region of NRF2 for GSK-3β.

Ser/The kinase	Rank by rate of phosphorylation	Relative activity to standard (activity compared as U/mg protein)	Peptide or protein standard	Relative phosphorylation to background (mouse/Human peptide)
DYRK2	1=	9	Peptide	50/10x
TTK	1=	40	Peptide	50/50x
MLK	1=	5	Protein	40/50x
Cdk9/CycT	4=	40	Peptide	20/10x
TBK1	4=	20	Peptide	20/15x
DYRK3	6=	25	Peptide	15/3x
CamKKb	7=	15	Peptide	12/10x
Wnk1	7=	15	Protein	12/12x
JNK1	7=	5	Protein	12/11x
SAPK2	10=	15	Protein	10/10x
NEK6	10=	13	Peptide	10/8x
SAPK4	10=	10	Protein	10/8x
STK33	10=	10	Peptide	10/10x
LKB1	14=	12	Peptide	8/9x
JNK2	15=	5	Protein	8/5x
JNK3	15=	2.5	Protein	8/7x
ERK5	17=	10	Protein	6/4x
PLK1	18=	20	Peptide	2.5/4x
IKKβ	18=	9/10	Peptide	4/2x
MLK3	20=	8	Protein	5/7x

Table 3.2.1: Kinase screen against mouse and human peptides indicates that several members of the CMGC family of kinases phosphorylate the Neh6 region

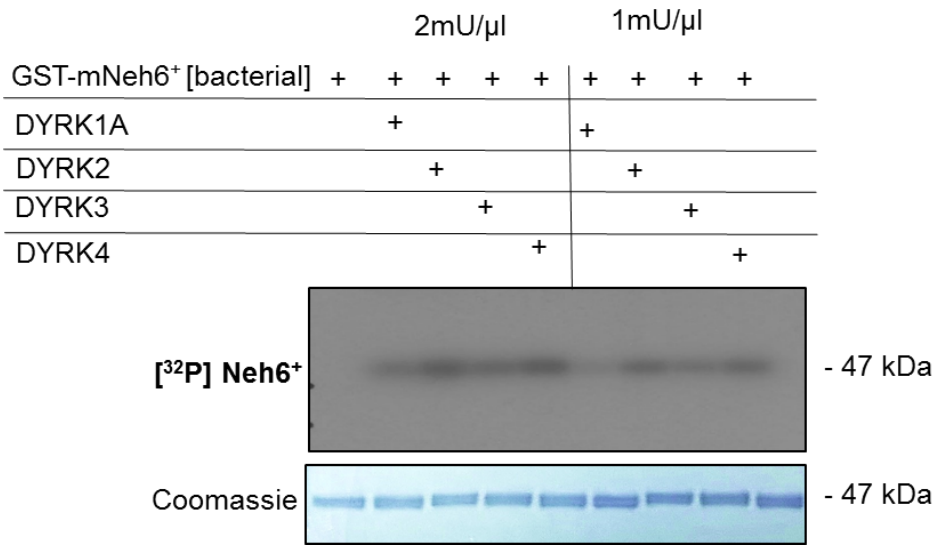
(A) A kinase screen was carried out using a mouse 21-mer peptide corresponding to a region in the Neh6 domain with the sequence FNDSDSGISLNTSPSRASPEH and a human 21-mer peptide with the sequence FNDSDSGISLNTSPSVASPEH; in both instances a KK dipeptide was attached to the N-terminal end to assist with immobilization of the peptides. The table depicts the top 20 kinases with the highest relative activity towards the mouse and human peptides compared to their standard peptides (positive controls). All of the top 20 kinases are members of the CMCG family of kinases. The kinases listed have been ranked in the second column according to their rate of phosphorylation of the 21-mer peptides. DYRK2, TTK and MLK1 were all ranked equal top.

3.2.4 All four DYRK family members phosphorylate the Neh6^{+} region of mNRF2 to varying abilities, with DYRK2 and DYRK4 showing the greatest activity

To determine the ability of the four DYRK family members to phosphorylate the Neh6 domain of mNRF2, we exploited the fact that bacterially expressed recombinant proteins are exposed to very little kinase activity and therefore will be unphosphorylated. Thus, GST-mNeh6^{+} was expressed and purified from *E.coli* (*BL21*(*DE3*)*pLysS*) cells and incubated with one of the four DYRK family members; DYRK1A, DYRK2, DYRK3 or DYRK4, in the presence of [γ-^{32}P]-ATP. It should be noted that there is an additional splice variant of DYRK1 called DYRK1B. However, DYRK1B has a very cell specific expression pattern and therefore was not included in our experiments.

DYRKs are a group of evolutionally conserved kinases that as their name suggests can phosphorylate both Tyr and Ser/Thr residues flanked by a C-terminal Pro residue. DYRKs are activated by autophosphorylation of a tyrosine residue located in their activation loop. DYRK family members have been reported to have roles in cell cycle, proliferation, mRNA splicing and synaptic signalling (Soppa and Becker, 2015). Before carrying out any experiments all kinases were assayed against a known substrate to determine their activity and ensure each kinase was used at the same activity (data not shown). The largest incorporation of phosphate into GST-mNeh6^{+} occurred in the presence of DYRK2 and DYRK4 suggesting they show greater activity towards the Neh6 domain than DYRK1A and DYRK3 (Figure 3.2.4). Phosphorylation of GST-mNeh6^{+} was greater when the kinases were at a concentration of 2 mU/μl than at 1 mU/μl, suggesting initial rate kinetics was reached.

A


2mU/µl
1mU/µl
GST-mNeh6+ [bacterial] + + + + + + + + + +
DYRK1A + +
DYRK2 + +
DYRK3 + +
DYRK4 + +
[32P] Neh6+
- 47 kDa
Coomassie
- 47 kDa

B

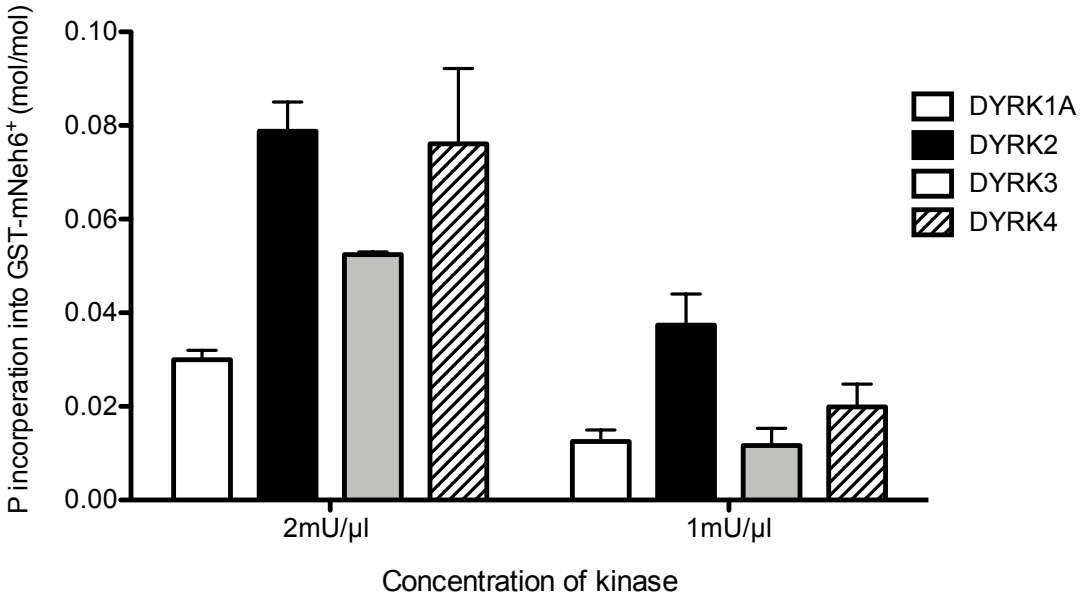

P incorperation into GST-mNeh6+ (mol/mol)
0.10
0.08
0.06
0.04
0.02
0.00
2mU/µl
1mU/µl
Concentration of kinase
DYRK1A
DYRK2
DYRK3
DYRK4

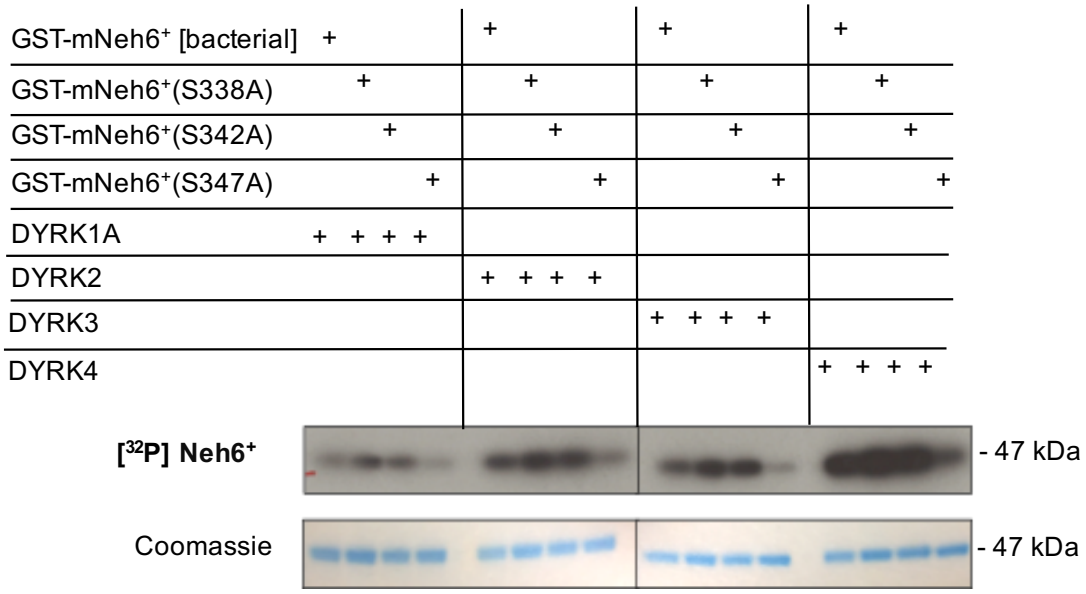
Figure 3.2.4: DYRK2 and DYRK4 phosphorylate the Neh6^{+} region of mNRF2 better than other DYRK family members

(A) To determine if all members of the DYRK family can phosphorylate the Neh6 domain of mNRF2, purified recombinant GST-tagged mNeh6^{+} protein (0.1 µg/µl, substrate) generated in E. coli (BL21(DE3)pLysS) cells was incubated (in a reaction mixture of 40 µl) with purified activated recombinant DYRK1A, DYRK2, DYRK3, or DYRK4, at a concentration of 2 mU/µl or 1mU/µl for 10 min in the presence of 10 mM $MgCl_2$ and 0.1 mM [γ-^{32}P]-ATP at 30°C. The reactions were terminated by addition of SDS gel loading buffer ,20 µl of each sample was loaded on a SDS-PAGE gel, and the excess [γ-^{32}P]-ATP removed by electrophoresis. Gels were then stained with coomassie dye, dried and subjected to autoradiography. After autoradiography, protein bands were excised from the gel and the incorporated radioactivity determined by scintillation counting. This image is one representative image of three independent biological replicates. (B) Scintillation counts from figure (A). DYRK1A is represented by the white bars, DYRK2 is shown by the black bars, DYRK3 is shown by the light grey bars and DYRK4 by the stripped bars. Kinase concentrations are given below the figure. The data shown is the average from three independent replicates and the associated SEM.

3.2.5 DYRK family members phosphorylate the Neh6 domain of mNRF2 predominantly at Ser-347

To identify which residue(s) is phosphorylated by the DYRK family of kinases, a series of point mutants (S338A, S342A and S347A) were generated in GST-mNeh6^{+} by mutating these Ser residues to an Ala, they can no longer be phosphorylated. GST-mNeh6^{+}, or the single base pair mutants generated were expressed and purified from *E.coli* (*BL21*(*DE3*)*pLysS*) cells, then incubated with; DYRK1A, DYRK2, DYRK3 or DYRK4. As shown in Figure 3.2.5, only when Ser-347 is mutated to Ala is there a loss of phosphorylation by all four DYRK kinases. This suggests that DYRK primes the Neh6 domain by phosphorylating Ser-347, independently of the phosphorylation status of neighbouring Ser residues. Interestingly, the DYRK isoenzmyes were all still able to phosphorylate the S342A mutant protein, suggesting this site is not a priming site but potentially a GSK-3β phosphorylation site.

A


GST-mNeh6+ [bacterial] + + + +
GST-mNeh6+(S338A) + + + +
GST-mNeh6+(S342A) + + + +
GST-mNeh6+(S347A) + + + +
DYRK1A + + + +
DYRK2 + + + +
DYRK3 + + + +
DYRK4 + + + +
[32P] Neh6+
- 47 kDa
Coomassie
- 47 kDa

B

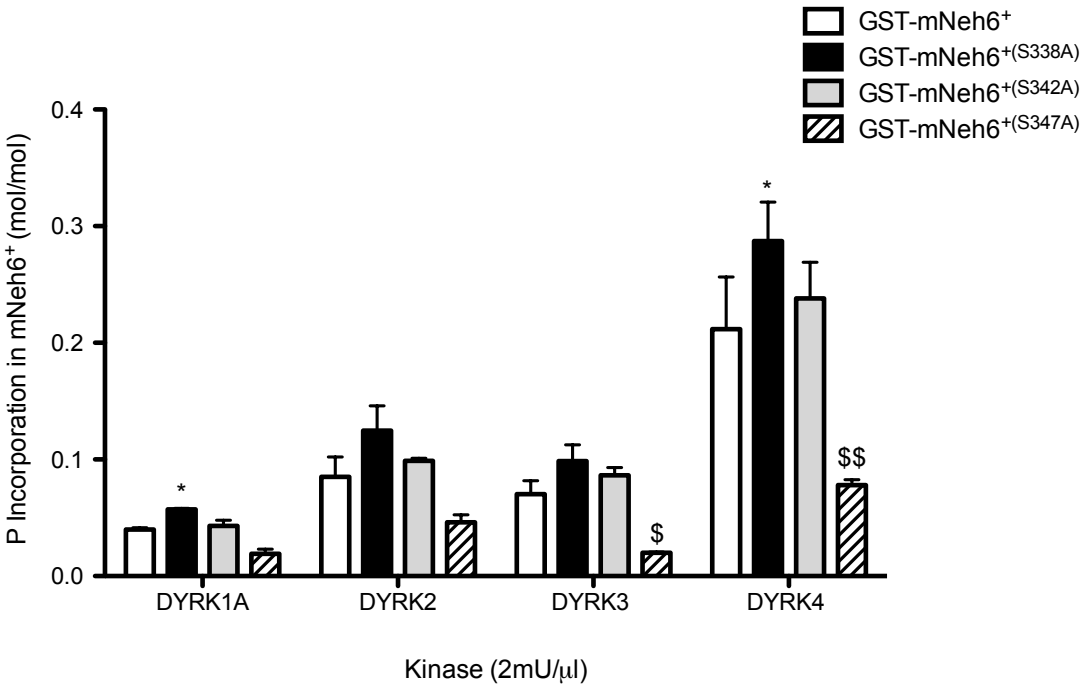

GST-mNeh6+
GST-mNeh6+(S338A)
GST-mNeh6+(S342A)
GST-mNeh6+(S347A)
P Incorporation in mNeh6+ (mol/mol)
0.4
0.3
0.2
0.1
0.0
*
*
$
$$
DYRK1A
DYRK2
DYRK3
DYRK4
Kinase (2mU/µl)

Figure 3.2.5: DYRK family members phosphorylate the Neh6 domain in mNRF2 predominantly at Ser-347

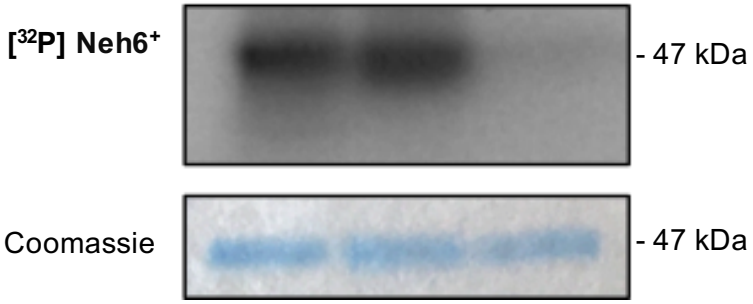
(A) To determine which Ser residue(s) in the Neh6 domain are targeted by members of the DYRK family of kinases, purified recombinant GST-tagged mNeh6^{+} protein or single base-pair mutants of mNeh6^{+} protein (0.1µg/µl, substrate) generated in E.coli (BL21(DE3)pLysS) cells were incubated in a reaction mixture of 40 µl with 2 mU/µl purified activated recombinant DYRK1A, DYRK2, DYRK3 or DYRK4, for 10 min in the presence of 10 mM $MgCl_2$ and 0.1 mM [γ-^{32}P]-ATP at 30°C. The reactions were terminated by addition of SDS gel loading buffer, then 20 µl of each sample was loaded on a SDS-PAGE gel, and the excess [γ-^{32}P]-ATP was removed by electrophoresis. Gels were then stained with coomassie dye, dried and subjected to autoradiography, after which, protein bands were excised and the incorporated radioactivity determined by scintillation counting. This image is representative of three independent biological replicates. (B) Scintillation counts from figure (A). GST-mNeh6^{+} is represented in the white bars, GST-mNeh6^{+}(S338A) is shown in the black bars, GST-mNeh6^{+}(S342A) is shown in the light grey bars and GST-mNeh6^{+}(S347A) in the striped grey bars. The identity of the kinases is given below the figure. The data displayed is the average from three independent replicates and the associated SEM.

3.2.6 DYRK2 phosphorylates Ser-347 enhancing subsequent phosphorylation of the Neh6 domain by GSK-3β

We next carried out a two-step linked assay with DYRK2 and GSK-3β to determine whether pre-incubating GST-mNeh6^{+} with DYRK2 leads to enhanced phosphorylation by GSK-3β. In the first step, bacterially purified GST-mNeh6^{+} or a single base-pair mutant protein, was incubated with DYRK2 in the presence of non-radioactively labelled ATP for 120 min; this first stage allows DYRK2 to phosphorylate GST-mNeh6^{+} priming it for GSK-3β. After 120 min, the reaction mix was heated at 60ºC for 5 min to inactivate the DYRK2 present. The reaction was then cooled before the addition of GSK-3β and [γ-^{32}P] ATP for 30 min. This means that the read out of this assay represents the ability of GSK-3β to phosphorylate GST-mNeh6^{+} protein. As shown in Figure 3.2.6, in the presence of DYRK2 and GSK-3β, GST-mNeh6^{+} is heavily phosphorylated. Mutation of Ser-342 to Ala does not affect GSK-3β phosphorylation, suggesting that DYRK2 phosphorylation is also unaltered. However, when Ser-347 is mutated to Ala there is a loss in GSK-3β phosphorylation, suggesting that priming by DYRK2 has been decreased.

A

GST-mNeh6^{+} [bacterial]	+	-	-
GST-mNeh6^{+}(S342A)	-	+	-
GST-mNeh6^{+}(S347A)	-	-	+
DYRK2	+	+	+
GSK-3β	+	+	+



B

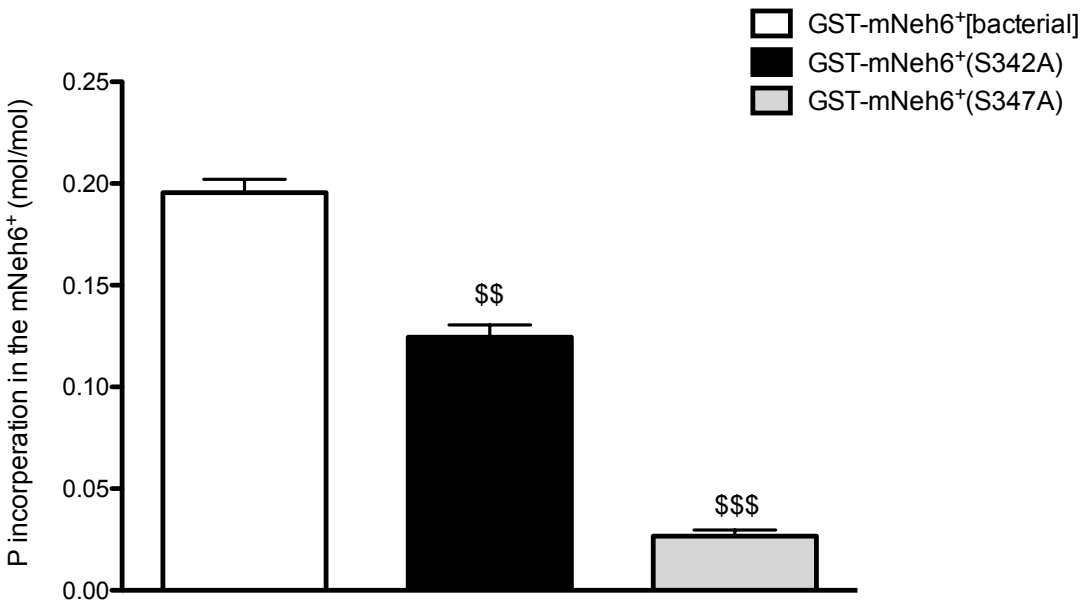


Figure 3.2.6: DYRK2 enhances phosphorylation of mNeh6^{+} by GSK-3β through phosphorylating S347A

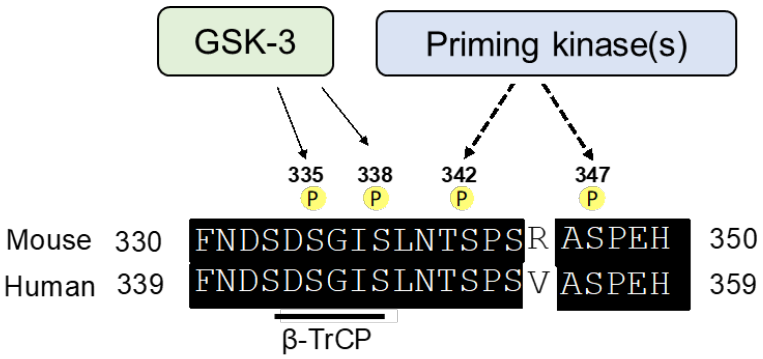
(A) A two-step in vitro kinase linking assay was setup to establish if DYRK2 can function as a priming kinase for GSK-3β by phosphorylating Ser-347. In the first step, purified recombinant GST-tagged mNeh6^{+} protein (0.1 µg/µl, substrate), or single base-pair mutant GST-mNeh6$^{+(S347A)}$ protein (0.1 µg/µl, substrate), generated in E.coli (BL21(DE3)pLysS) cells were incubated (in a reaction mixture of 4 0µl) with 2 mU/µl purified activated recombinant DYRK2, in the presence of 5 mM $MgCl_2$ and 0.05 mM cold-ATP at 30°C for 120 min. The reaction mixture was heated at 60°C for 5 min to inactivate the kinase and re-incubated at 30°C for 10 min before initiating the second step of the two-step of the assay. Purified activated recombinant GSK-3β (6 mU/µl (kinase)) was then added to the reaction mixtures in the presence of 5 mM $MgCl_2$ and 0.05 mM [γ-^{32}P]-ATP and incubated at 30°C for 60 min. Thereafter, the reactions were terminated by addition of SDS gel loading buffer, 20 µl of each sample was loaded onto a SDS-PAGE gel, and the excess [γ-^{32}P]-ATP was removed by electrophoresis. Gels were then stained with coomassie dye, dried and subjected to autoradiography, after which protein bands were excised and the incorporated radioactivity determine by scintillation counting. The image is from one experiment and is representative of three independent biological replicates. (B) Scintillation counts for (A). GST-mNeh6^{+} is represented in the white bar, GST-mNeh6^{+}(S342A) is shown in the black bar and GST-mNeh6^{+} is shown in the light grey bar. The data displayed is the average from three independent replicates and the associated SEM.

3.2.7 The identity of the amino acid residue at position 345 influences the ability of DYRK2 to phosphorylate the Neh6 domain in mNRF2

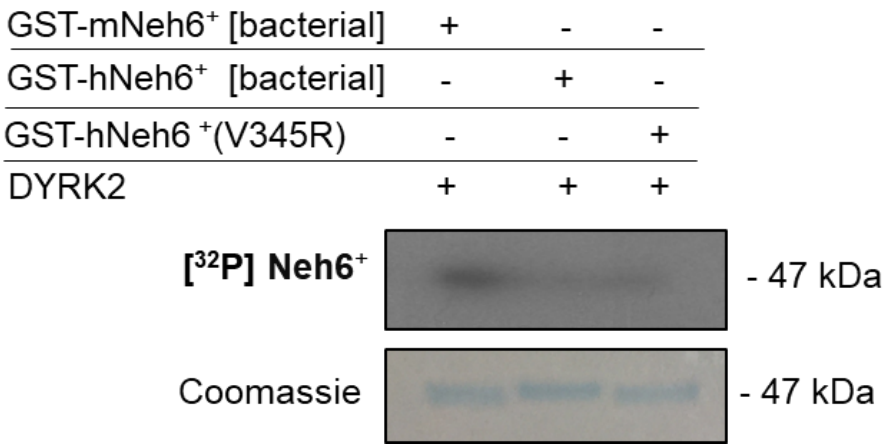
The *in vitro* kinase screen (Table 3.2.1) revealed that both DYRK2 and DYRK3 can phosphorylate the 21-mer peptide designed from mNRF2 considerably better than the equivalent human 21-mer peptide. Interestingly, only one amino acid difference exists between the mouse and the human peptide sequences; in mNRF2, the amino acid at position 345 is an Arg (R) residue, whereas the corresponding amino acid in hNRF2 is a Val (V) residue. To determine whether the identity of the amino acid at this position influences phosphorylation by DYRK2, we generated a fusion protein compromising GST linked at its C-terminal end to amino acids 299-419 of hNRF2, which we refer to as GST-hNeh6^{+}, and single base-pair V345R mutant. We then incubated bacterially purified GST-mNeh6^{+}, GST-hNeh6^{+} and GST-hNeh6^{+}(V345R) with DYRK2 in the presence of [γ-^{32}P] ATP (Figure 3.2.7). Phosphorylation of GST-

hNeh6^{+} by DYRK2 was considerably weaker than phosphorylation of GST-mNeh6^{+}, matching up with the data we saw in the kinase screen. Remarkably, when Val-354 in GST-hNeh6^{+} was replaced with a Arg residue, phosphorylation by DYRK2 was increased, though not to the extent observed with GST-mNeh6^{+} as substrate. These findings suggest that the identity of the amino acid positioned two residues N-terminal to the DYRK2 phosphorylation site can affect its activity towards the substrate.

A



B



C

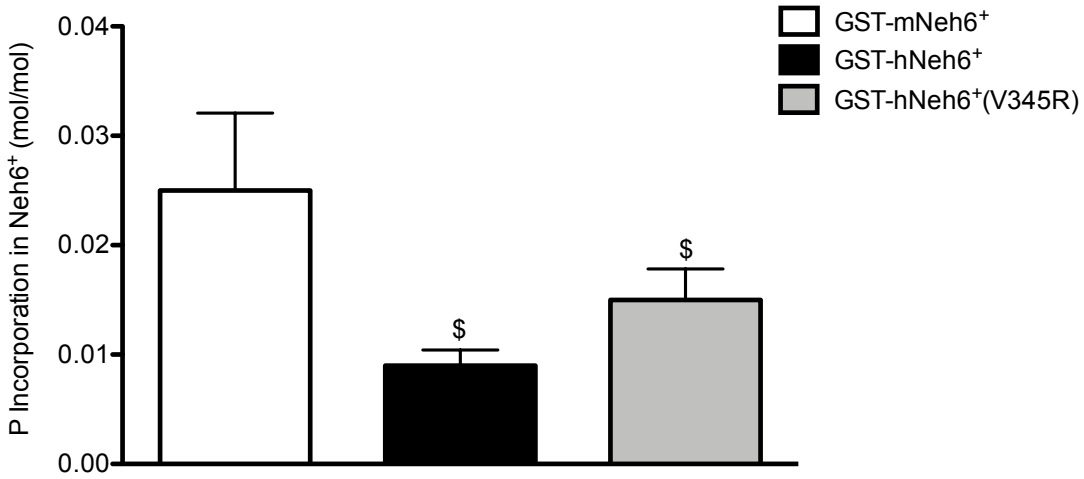


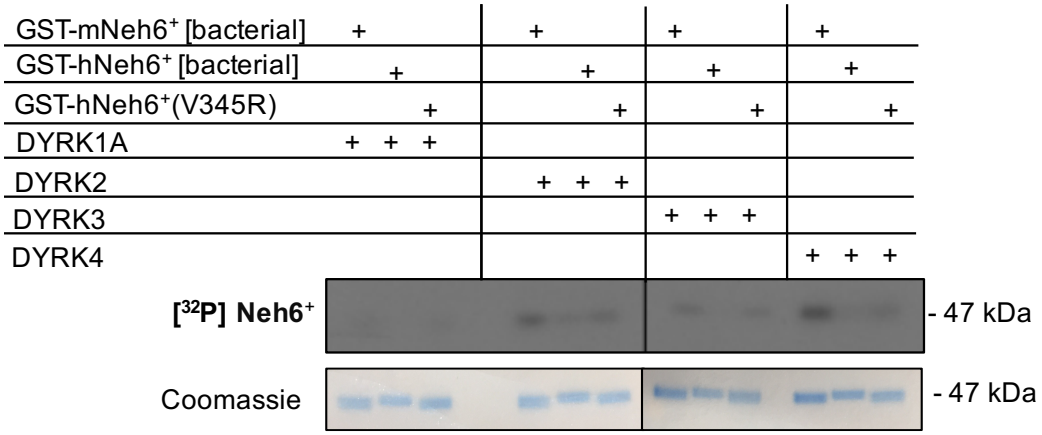
Figure 3.2.7: The identity of the amino acid at position 345 influences phosphorylation of the Neh6 domain by DYRK2

(A) Sequence alignment of the region surrounding the GSK-3 phosphorylation sites in mouse and hNRF2 reveals only one amino acid difference between the two species; this is Arg-345 in mNRF2, which is equivalent to Val-354 in hNRF2. (B) To determine if the identity of amino acid 345 (mouse numbering) influences DYRK2 phosphorylation of the Neh6 domain, purified recombinant GST-tagged mNeh6^{+} protein (0.1 µg/µl, substrate) or purified recombinant GST-hNeh6^{+} protein (0.1 µg/µl, substrate) or a single site mutant of GST-hNeh6^{+}(V345R) protein (0.1 µg/µl, substrate) generated in E.coli (BL21(DE3)pLysS) cells were incubated in a reaction mixture of 40 µl with purified activated recombinant DYRK2 (2 mU/µl) for 10 min in the presence of 10 mM $MgCl_2$ and 0.1 mM [γ-^{32}P]-ATP at 30°C. The reactions were terminated by addition of SDS gel loading buffer, the samples (20 µl) were loaded onto a SDS-PAGE gel, and the excess [γ-^{32}P]-ATP was removed by electrophoresis. Gels were then stained with Coomasssie Brilliant Blue dye, dried and subjected to autoradiography, after which protein bands were excised and the incorporated radioactivity determined by scintillation counting. This image is from one experiment and is representative of three independent biological replicates. (C) Scintillation counts obtained from (B). GST-mNeh6^{+} is represented in the white bar, GST-hNeh6^{+} is shown in the black bar and GST-hNeh6^{+}(V345R) in the light grey bar. The data displayed is the average from three independent replicates with the associated SEM.

3.2.8 The ability of all DYRK family members to phosphorylate the Neh6 domain is influenced by the identify of amino acid 345

To establish whether the identify of amino acid 345 influences the ability of all four DYRK family members to phosphorylate the Neh6 domain, we carried out an *in vitro* kinase assay with all four isoenzymes in the presence of bacterially purified GST-mNeh6$^+$, GST-hNeh6$^+$ and GST-hNeh6$^+$(V345R). All four DYRK family members were able to phosphorylate GST-mNeh6$^+$ to a greater extent than GST-hNeh6$^+$ (Figure 3.2.8). As seen in the previous figure, mutating Val-345 in GST-hNeh6$^+$ to an Arg, partially rescued phosphorylation. This reinforces the idea that the identity of neighbouring residues around the phosphorylation site influences the ability of DYRK family members to phosphorylate mNRF2.

A



B

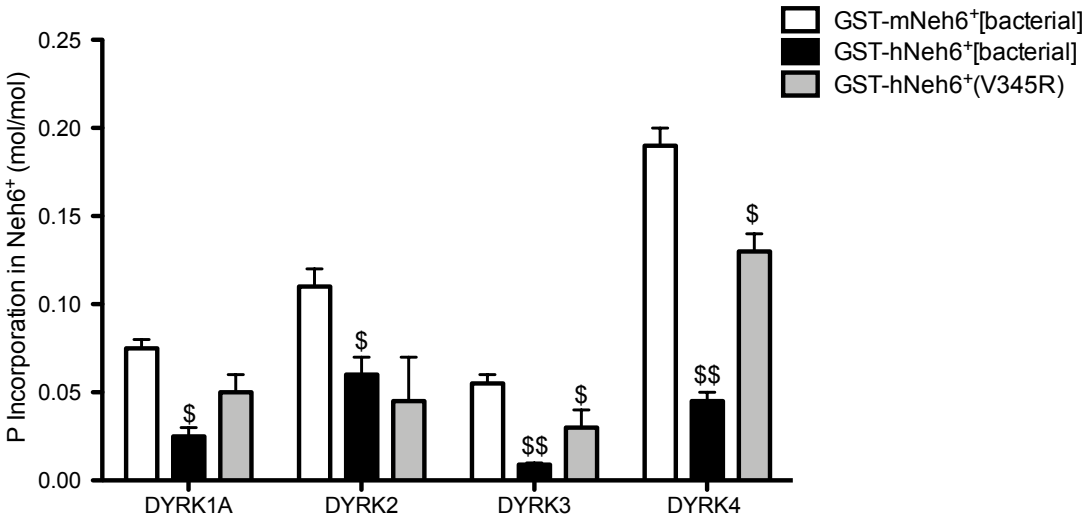


Figure 3.2.8: The identity of the amino acid at position 345 influences the ability of all DYRK isoenzymes to phosphorylate the Neh6 domain.

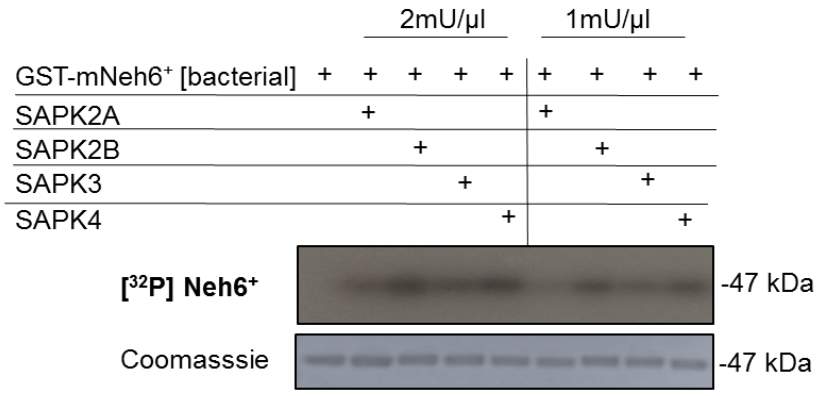
(A)To determine if amino acid 345 (mouse numbering) influences the ability of all four DYRK family members to phosphorylate the Neh6 domain, purified recombinant GST-mNeh6^{+} protein (0.1 µg/µl) or purified recombinant GST-hNeh6^{+} protein (0.1 µg/µl) or a single site mutant of GST-hNeh6^{+}(V345R) protein (0.1 µg/µl) generated from E.coli (BL21(DE3)pLysS) cells were incubated in a total reaction mix of 40 µl with purified activated recombinant DYRK1A, DYRK2, DYRK3 or DYRK4 (each at 2 mU/µl) for 10 min in the presence of 10 mM $MgCl_2$ and 0.1 mM [γ-^{32}P]-ATP at 30°C. The reactions were terminated by addition of SDS gel loading buffer, 20 µl of each sample was then loaded on a SDS-PAGE gel, and the excess [γ-^{32}P]-ATP was removed by electrophoresis. Gels were then stained with coomasssie dye, dried and subjected to autoradiography, after which protein bands were excised and the incorporated radioactivity determined by scintillation counting. The data shown are from one experiment and is representative of three independent biological replicates. (B) Scintillation counts from (A).
GST-mNeh6^{+} is represented in the white bars, GST-hNeh6^{+} is shown in the light grey bars and GST-hNeh6^{+}(V345R) in the light grey bars. The identity of the kinases is given below the figure. The data shown is the average of three independent replicates with the associated SEM.

3.2.9 Kinases identified in the 21-mer peptide screen phosphorylate the GST-mNeh6^{+} fusion protein through different sites

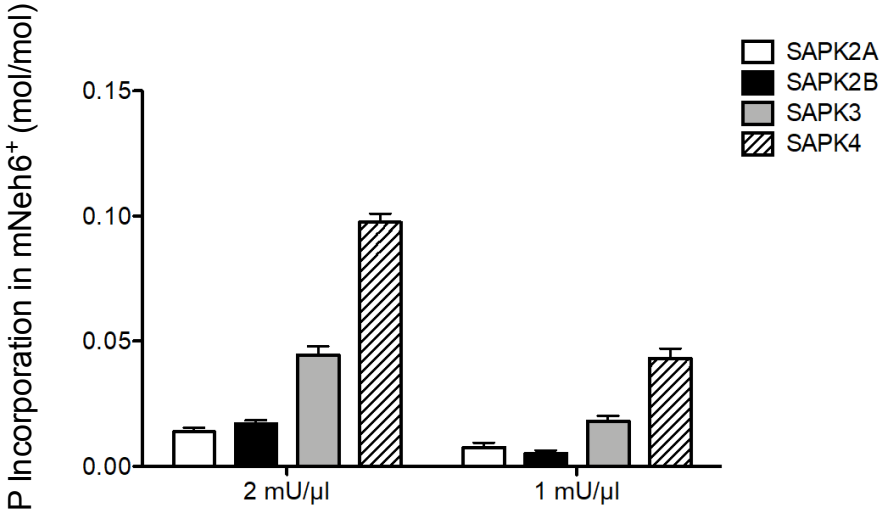
To further validate the kinase screen data (Table 3.2.1), we carried out several *in vitro* kinase assays against purified bacterially-synthesised GST-mNeh6^{+}. As shown in Figure 3.2.9A and B, all four SAPK family members are able to phosphorylate GST-mNeh6^{+}, to varying extents, with SAPK3 and SAPK4 exhibiting the highest phosphorylation. To determine whether this phosphorylation of the mNeh6^{+} domain by SAPK isoenzymes occurs through at Ser-347, the same site as utilized by DYRK isoenzymes, or thorough the other potential Ser-342 priming site, we incubated bacterially purified GST-mNeh6^{+}, or single amino acid site mutant constructs with SAPK2A, SAPK2B, SAPK3 or SAPK4. By contrast with data obtained with DYRK isoenzymes (Figure 3.2.5), mutation of Ser-347 to Ala had little effect on the ability of SAPK isoenzymes to phosphorylate GST-mNeh6^{+} (Figure 3.2.9C and D). However, replacement of Ser-342 with an Ala residue decreased the ability of all four SAPK isoenzymes to phosphorylate the mutant GST-mNeh6^{+} protein, suggesting that this family of

protein kinases may prime mNRF2 for modification by GSK-3β through phosphorylation of Ser-342.

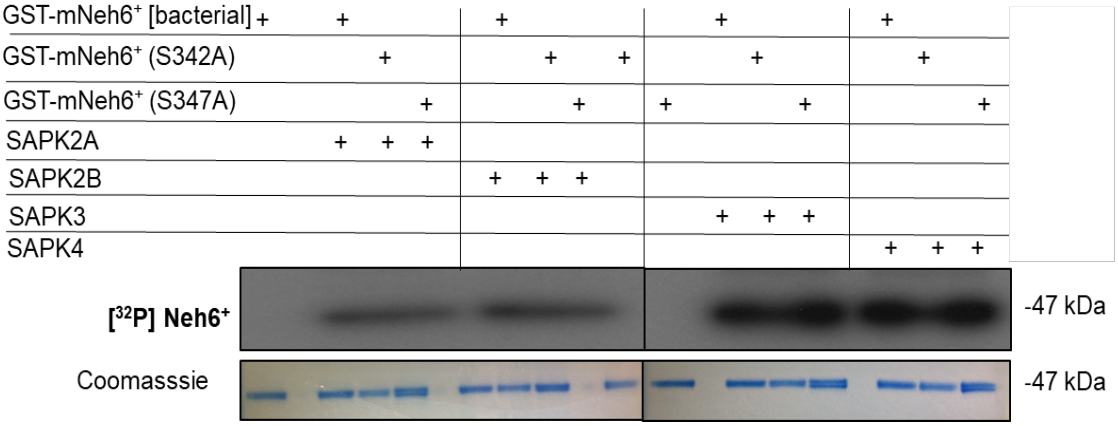
A



B



C


GST-mNeh6+ [bacterial]
GST-mNeh6+ (S342A)
GST-mNeh6+ (S347A)
SAPK2A
SAPK2B
SAPK3
SAPK4
[32P] Neh6+
-47 kDa
Coomasssie
-47 kDa

D

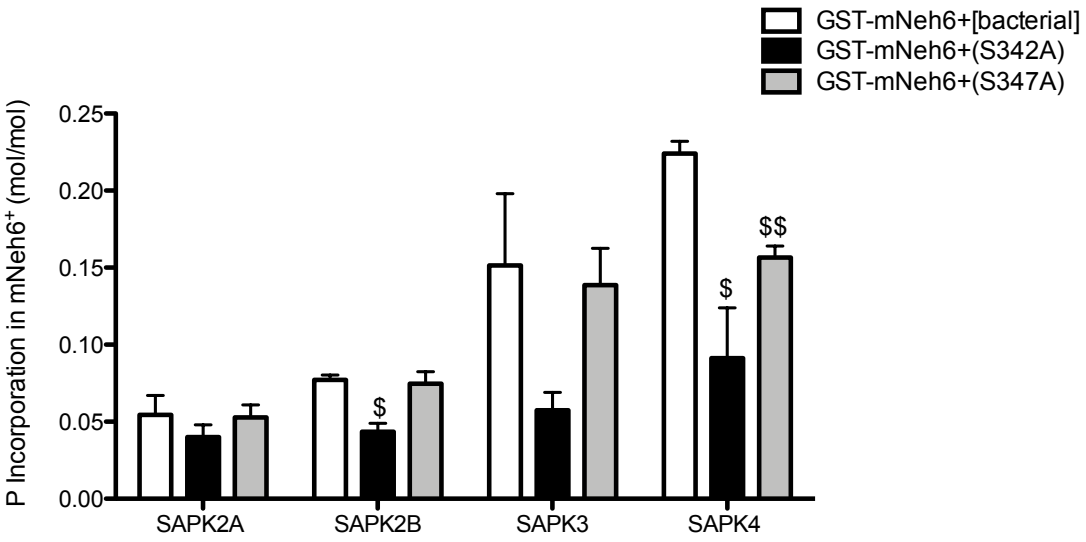
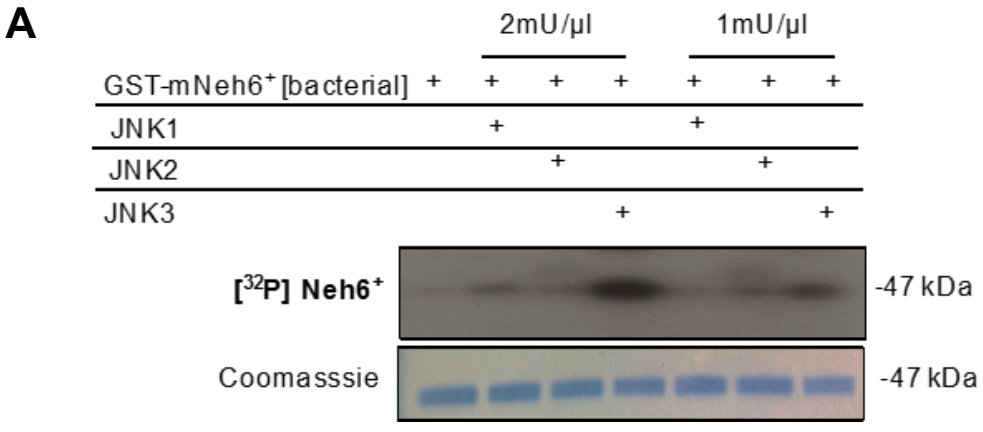

GST-mNeh6+[bacterial]
GST-mNeh6+(S342A)
GST-mNeh6+(S347A)
P Incorporation in mNeh6+ (mol/mol)
0.25
0.20
0.15
0.10
0.05
0.00
$
$
$$
SAPK2A
SAPK2B
SAPK3
SAPK4

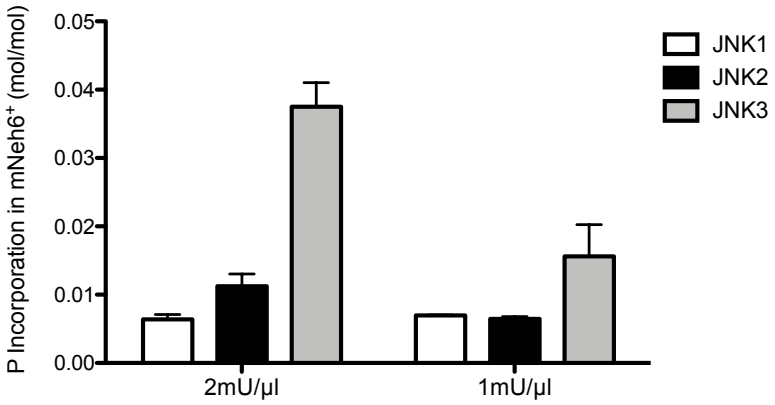
Figure 3.2.9: Members of the SAPK family phosphorylate the Neh6 domain of mNRF2 predominantly through Ser-342.

(A) To determine if SAPK family members can phosphorylate the Neh6 domain of mNRF2, purified recombinant GST-tagged mNeh6^{+} protein (0.1 µg/µl, substrate) generated in E.coli (BL21(DE3)pLysS) cells was isolated on glutathione-agarose beads and incubated with purified activated recombinant SAPK2A, SAPK2B, SAPK3, or SAPK4, at either at concentration of 2 mU/µl or 1 mU/µl in a reaction mixture of 40 µl for 10 min in the presence of 10 mM $MgCl_2$ and 0.1 mM [γ-^{32}P]-ATP at 30°C. The reactions were terminated by addition of SDS gel loading buffer, 20 µl of each sample was then loaded on a SDS-PAGE gel, and excess [γ-^{32}P]-ATP was removed by electrophoresis. Gels were then stained with Coomassie Brilliant Blue dye, dried, and subjected to autoradiography, after which protein bands were excised and the incorporated radioactivity determined by scintillation counting. This image is one representative image of three independent biological replicates. (B) Scintillation counts from (A) SAPK2A is represented by the white bars, SAPK2B is shown in the black bars, SAPK3 is shown in the light grey bars and SAPK4 is shown in the stripped bars. Kinase concentrations are given below the figure. This data shown is the mean value from three replicates and the associated SEM.(C) To determine which Ser residue in the Neh6 domain is targeted by members of the SAPK family, purified recombinant GST-mNeh6^{+} protein or single base-pair mutants for mNeh6^{+} protein (0.1 µg/µl, substrate) generated in E.coli (BL21(DE3)pLysS) cells were incubated in a reaction mixture of 40 µl with 2 mU/µl purified activated recombinant SAPK2A, SAPK2B, SAPK3 or SAPK4, for 10 min in the presence of 10 mM $MgCl_2$ and 0.1 mM [γ-^{32}P]-ATP at 30°C. The reactions were terminated by addition of SDS gel loading buffer, 20 µl of each sample was then loaded on a SDS-PAGE gel, and the excess [γ-^{32}P]-ATP was removed by electrophoresis. Gels were then stained with coomassie dye, dried, and subjected to autoradiography, after which, protein bands were excised and the incorporated radioactivity determined by scintillation counting. This image is representative of three independent biological replicates. (D) Scintillation counts from (C). GST-mNeh6^{+} is represented in the white bars, GST-mNeh6^{+}(S342A) is shown in the black bars and GST-mNeh6^{+}(S347A) is shown in the light grey bars. The identity of the kinase is given below the figure. The data displayed is the mean value from three independent replicates and the associated SEM.

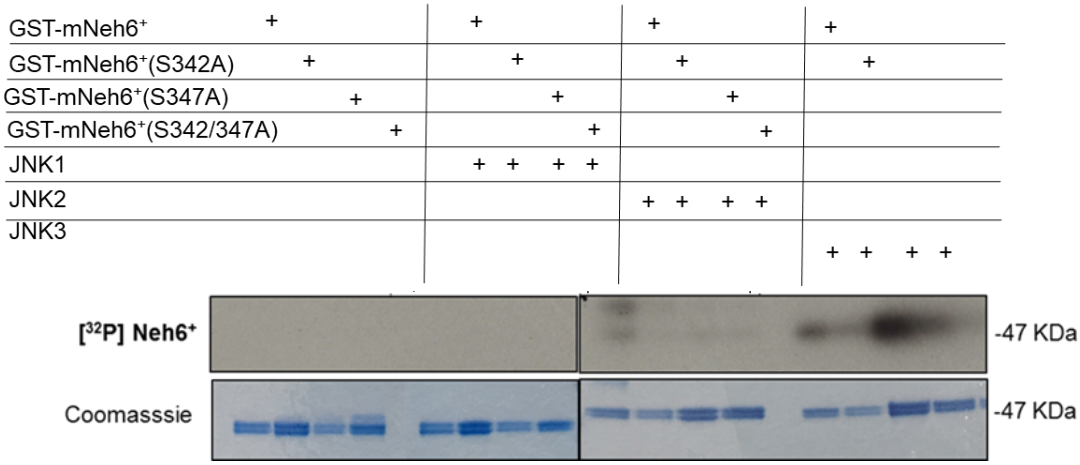
Of the kinases examined, all three JNK family members: JNK1, JNK2 and JNK3, were amongst the most active in phosphorylating the 21-mer peptides in the *in vitro* screen (Table 3.2.1). To determine whether all three JNK isoenzymes phosphorylate the Neh6 domain to a similar extent, we incubated bacterially synthesised GST-mNeh6^{+} with either JNK1, JNK2 or JNK3. As seen in Figure 3.2.10A and B, JNK3 seems better at phosphorylating GST-mNeh6^{+} than the two other JNK family members. This was surprising as the screen results suggest that JNK1 should have been better than JNK2 and JNK3 at phosphorylating Neh6. Potentially there is altered folding in the larger fusion protein than the 21-mer peptide used in the screen, which may alter the ability of some kinases to bind. To investigate whether phosphorylation of the Neh6 domain by JNK isoenzymes occurs through either Ser-342 or Ser-347, bacterially purified GST-mNeh6^{+}, single amino acid site mutants of GST-mNeh6^{+} or a double mutant of GST-mNeh6^{+}, in which both potential priming sites are mutated, were incubated with the three JNK proteins. The data displayed in Figure 3.2.10C and D show that phosphorylation by all three JNK family members is reduced in the fusion protein containing S342A or in the S342A/S347A double mutant protein. Phosphorylation by JNK3 appears to be enhanced when Ser-347 is mutated to Ala. Taking all the data in Figure 3.2.10 together, it is apparent that JNK isoenzymes can phosphorylate mNeh6^{+} but it is currently unclear which site they target.



B


P Incorporation in mNeh6+ (mol/mol)
0.05
0.04
0.03
0.02
0.01
0.00
2mU/µl
1mU/µl
JNK1
JNK2
JNK3

C


GST-mNeh6+
GST-mNeh6+(S342A)
GST-mNeh6+(S347A)
GST-mNeh6+(S342/347A)
JNK1
JNK2
JNK3
[32P] Neh6+
-47 KDa
Coomasssie
-47 KDa

D

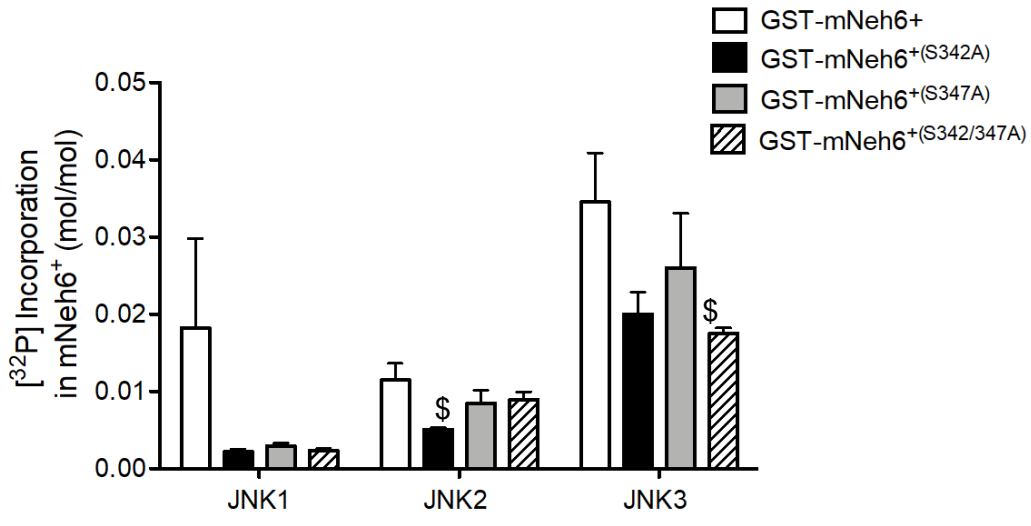
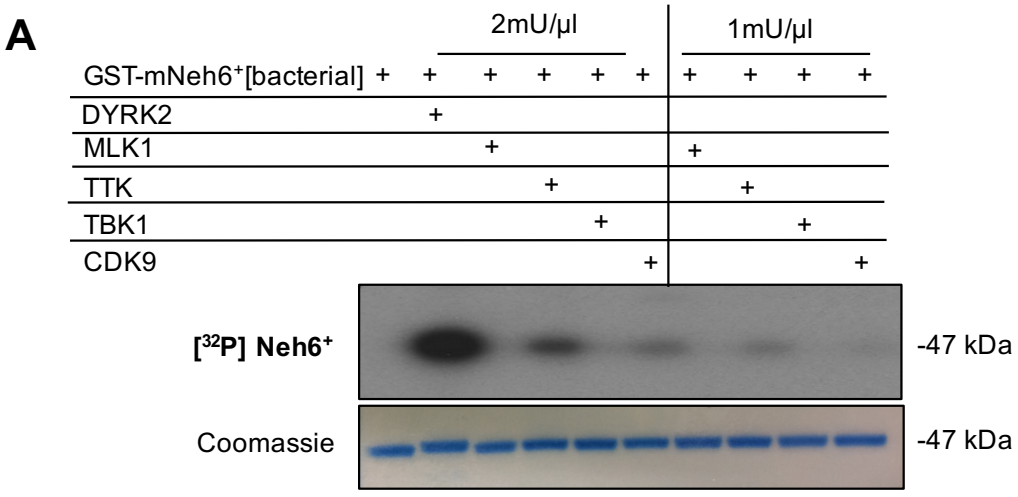

GST-mNeh6+
GST-mNeh6+(S342A)
GST-mNeh6+(S347A)
GST-mNeh6+(S342/347A)
[32P] Incorporation in mNeh6+ (mol/mol)
0.05
0.04
0.03
0.02
0.01
0.00
$
$
JNK1
JNK2
JNK3

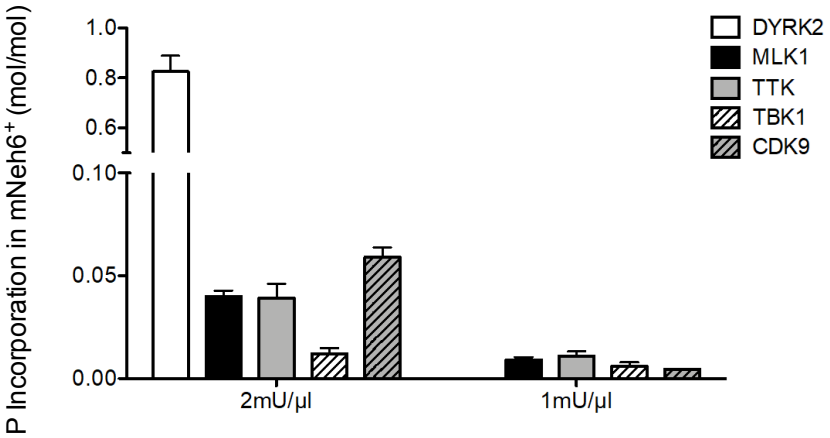
Figure 3.2.10: **Members of the JNK family can also phosphorylate the Neh6 region, but it is unclear through which site**

(A) To determine if different members of the JNK family can phosphorylate the Neh6 region to different abilities, purified recombinant GST-tagged mNeh6^{+} protein (0.1 µg/µl, substrate) generated from E.coli (BL21(DE3)pLysS) cells was incubated with purified activated recombinant: JNK1, JNK2, or JNK3, at either at concentration of 2 mU/µl or 1 mU/µl for 30 min in the presence of 10 mM $MgCl_2$ and 0.1 mM [ɣ-^{32}P]-ATP at 30°C. The reactions were terminated by addition of SDS gel loading buffer, the samples were loaded on a SDS-PAGE gel, and the excess [ɣ-^{32}P]-ATP was removed by electrophoresis. Gels were then stained with coomasssie, dried and subjected to autoradiography. After which, protein bands were excised and the incorporated radioactivity was determined by scintillation counting. This image is one representative image of three independent biological replicates. (B) Scintillation counts from figure (A). JNK1 is represented in the white bars, JNK2 is shown in the black bars and JNK3 is shown in the light grey bars. Kinase concentrations are displayed under the figure. The data displayed on the graph is the mean value from three independent biological replicates and the associated SEM. (C) To determine which serine residue in the Neh6 region is targeted by members of the JNK family of kinases, purified recombinant GST-tagged mNeh6^{+} protein or single/double base-pair mutants for mNeh6^{+} protein (0.1 µg/µl, substrate) generated from E.coli (BL21(DE3)pLysS) cells were incubated with 2 mU/µl purified activated recombinant, JNK1, JNK2 or JNK3, for 30 minutes in the presence of 10 mM $MgCl_2$ and 0.1 mM [ɣ-^{32}P]-ATP at 30°C. The reactions were terminated by addition of SDS gel loading buffer, the samples were loaded on a SDS-PAGE gel, and the excess [ɣ-^{32}P]-ATP was removed by electrophoresis. Gels were then stained with coomasssie, dried and subjected to autoradiography. After which, protein bands were excised and the incorporated radioactivity was determined by scintillation counting. This image is representative of three independent biological replicates. (D) Scintillation counts from figure (C). GST-mNeh6^{+} is represented by the white bars, GST-mNeh6^{+}(S342A) is shown in the black bars, GST-mNeh6^{+} (S347A) is shown in the light grey bars and GST-mNeh6^{+}(S342/347A) is shown in the stripped bars. The identity of the kinase is shown below the figure. The data displayed on the graph is the mean value from three independent biological replicates and the associated SEM.

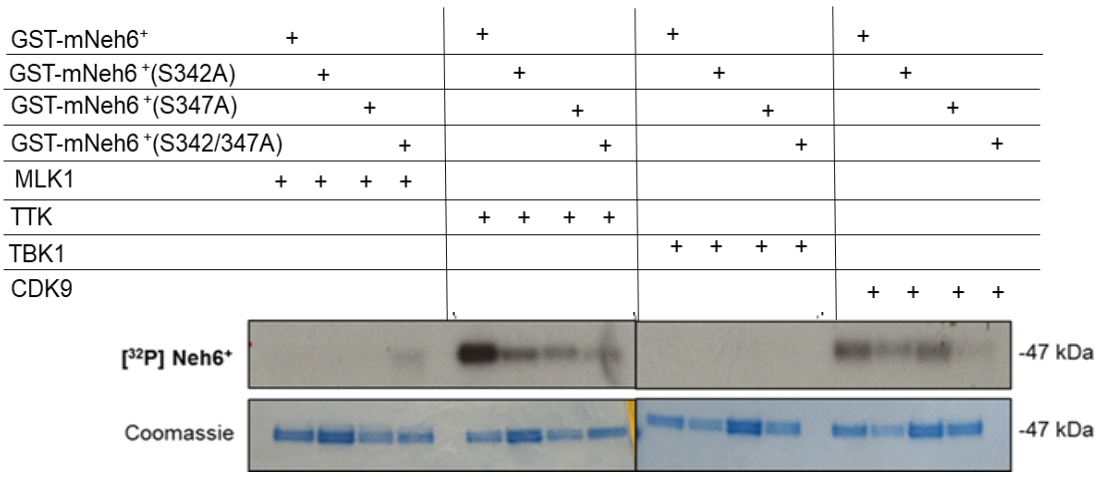
Using the 21-mer peptide-based screen, TTK and MLK1 came out ranked equal first with DYRK2, and CDK9 and TBK1 were ranked equal fourth. However, when we compared the ability of TTK, MLK1, CDK9 and TBK1 to phosphorylate bacterially synthesised GST-mNeh6^{+} purified on glutathione-agarose beads to that of DYRK2, we found that all four kinases showed significantly less phosphorylation of the fusion protein relative to DYRK2, as shown is Figure 3.2.11A and B. To determine which specific site these four kinases phosphorylate in the Neh6 domain, we incubated bacterially purified GST-mNeh6^{+}, or a single amino acid site mutant or a double mutant protein with either MLK1, TTK, TBK1 or CDK9. As seen is Figure 3.2.11C and D, mutation of Ser-342 and Ser-347 lead to a reduction in phosphorylation by TTK. Whilst mutation of Ser-347 to Ala did not influence phosphorylation by CDK9 and mutation of Ser-342 to Ala lead to a slight reduction. Phosphorylation of any of the fusion proteins by either MLK1 and TBK1 was too low to draw any conclusions.



B


P Incorporation in mNeh6+ (mol/mol)
1.0
0.8
0.6
0.10
0.05
0.00
2mU/µl
1mU/µl
DYRK2
MLK1
TTK
TBK1
CDK9

C


GST-mNeh6+
GST-mNeh6+(S342A)
GST-mNeh6+(S347A)
GST-mNeh6+(S342/347A)
MLK1
TTK
TBK1
CDK9
[32P] Neh6+
-47 kDa
Coomassie
-47 kDa

D

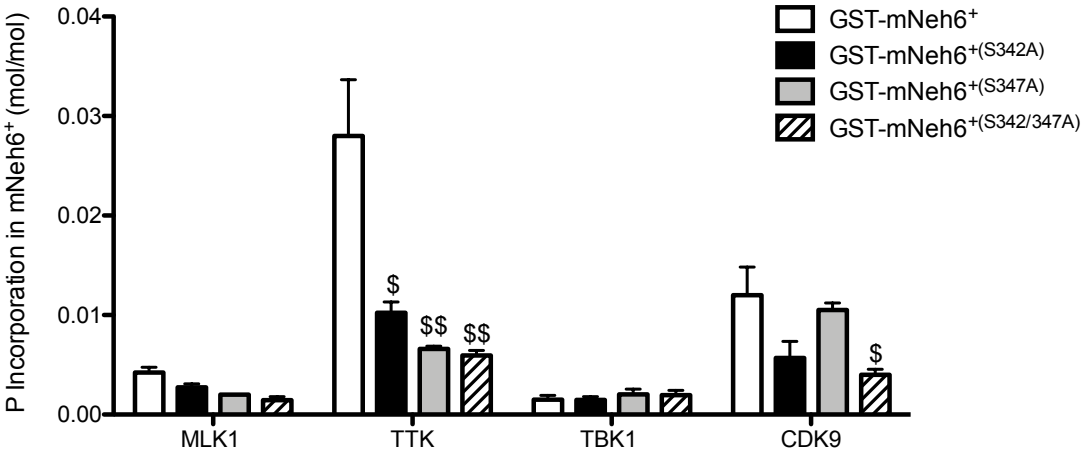

P Incorporation in mNeh6+ (mol/mol)
0.04
0.03
0.02
0.01
0.00
$
$$
$$
$
MLK1
TTK
TBK1
CDK9
GST-mNeh6+
GST-mNeh6+(S342A)
GST-mNeh6+(S347A)
GST-mNeh6+(S342/347A)

Figure 3.2.11*: *MLK1, TTK, TBK1 and CDK9 phosphorylate the Neh6 domain to varying degrees

(A) To determine if MLK1, TTK, TBK1 and CDK9, can phosphorylate the Neh6 region, purified recombinant GST-tagged mNeh6^{+} protein (0.1 µg/µl) generated in E.coli (BL21(DE3)pLysS) cells was incubated in a reaction mixture of 40 µl with purified activated recombinant MLK1, TTK, TBK1 or CDK9, at either at concentration of 2 mU/µl or 1 mU/µl for 3 hr in the presence of 10 mM $MgCl_2$ and 0.1 mM [γ-^{32}P]-ATP at 30°C; purified recombinant DYRK2 (2 mU/µl) was used as a positive control. The reactions were terminated by addition of SDS gel loading buffer, 20 µl of each sample was loaded on a SDS-PAGE gel, and the excess [γ-^{32}P]-ATP removed by electrophoresis. Gels were then stained with Coomassie dye, dried and subjected to autoradiography, after which protein bands were excised and the incorporated radioactivity determined by scintillation counting. The image is representative of three independent biological replicates. (B) Scintillation counts from (A). DYRK2 is represented by the white bars, MLK1 is shown in the black bars, TTK is shown in the light grey bars, TBK1 is shown in the stripped bar and CDK9 is shown in the stripped grey bar. Kinases concentrations are indicated below the figure. The data displayed is the average from three independent replicates and the associated SEM. (C) To determine which Ser residue in the Neh6 domain is targeted by MLK1, TTK, TBK1 and CDK9, purified recombinant GST-tagged mNeh6^{+} protein or single/double base-pair mutants of the mNeh6^{+} protein (0.1 µg/µl) generated from E.coli (BL21(DE3)pLysS) cells were incubated in a reaction mixture of 40 µl with 2 mU/µl purified activated recombinant MLK1, TTK,TBK1 or CDK9, for 3 hr in the presence of 10 mM $MgCl_2$ and 0.1 mM [γ-^{32}P]-ATP at 30°C. The reactions were terminated by addition of SDS gel loading buffer, 20 µl of each sample was then loaded on a SDS-PAGE gel, and the excess [γ-^{32}P]-ATP removed by electrophoresis. Gels were then stained with coomassie dye, dried and subjected to autoradiography, after which, protein bands were excised and the incorporated radioactivity determined by scintillation counting. This image is representative of three independent biological replicates. (D) Scintillation counts from (C). GST-mNeh6^{+} is represented in the white bars, GST-mNeh6^{+}(S342A) is shown in the black bars, GST-mNeh6^{+}(S347A) is shown in the light grey bar and GST-mNeh6^{+}(S342/347A) is shown in the stripped bar. The identity of kinase is given below the figure. The data displayed is the average from three independent replicates and the associated SEM.

3.2.10 Immunochemical validation that DYRK2 and SAPK3 phosphorylate the mNeh6 domain at Ser-342 and Ser-347, respectively

To provide further evidence that DYRK2 and SAPK3 phosphorylate the Neh6 domain of mNRF2 through Ser-342 and Ser-347 respectively, an immunochemical-based assay was carried out using our in-house phospho-specific antibodies raised against peptides that contained a phosphorylated Ser residue at positions equivalent to either Ser-342 or Ser-347. In these experiments, bacterially synthesised GST-mNeh6^{+}, a single amino acid site mutant or a double amino acid mutant protein was incubated with either DYRK2 or SAPK3 for 1 hr in the presence of non-radioactively labelled ATP. Samples were then loaded on to a SDS-PAGE gel, transferred to PVDF membrane and blocked in 5% (w/v) milk for 1 hr before probing overnight with the appropriate phospho-specific antibody; all phospho-specific antibodies underwent dot blotting to determine their specificity and dilution before use, as shown in Appendix 8.6. No phosphorylation was detected by the P-338 (which detects phosphorylation at Ser-338) NRF2 antibody in any of the experimental conditions suggesting that is site is only phosphorylated in the presence of GSK-3β (Figure 3.2.12). Phosphorylation of Ser-342 appears to be still slightly detected by the P-342 (which detects phosphorylation at Ser-342) NRF2 antibody when the site is mutated to Ala indicating the antibody may not by phospho-specific for this site. Detected phosphorylation of Ser-342 is enhanced in the presence of SAPK3 but not DYRK2, and the opposite is seen in terms of phosphorylation of Ser-347. Only in the presence of DYRK2 is phosphorylation at Ser-347 detected, and in the presence of GST-mNeh6^{+}(S347A) or GST-mNeh6^{+}(S342/347A) proteins the phosphorylation at this site is completely lost suggesting the antibody is highly phospho-specific.

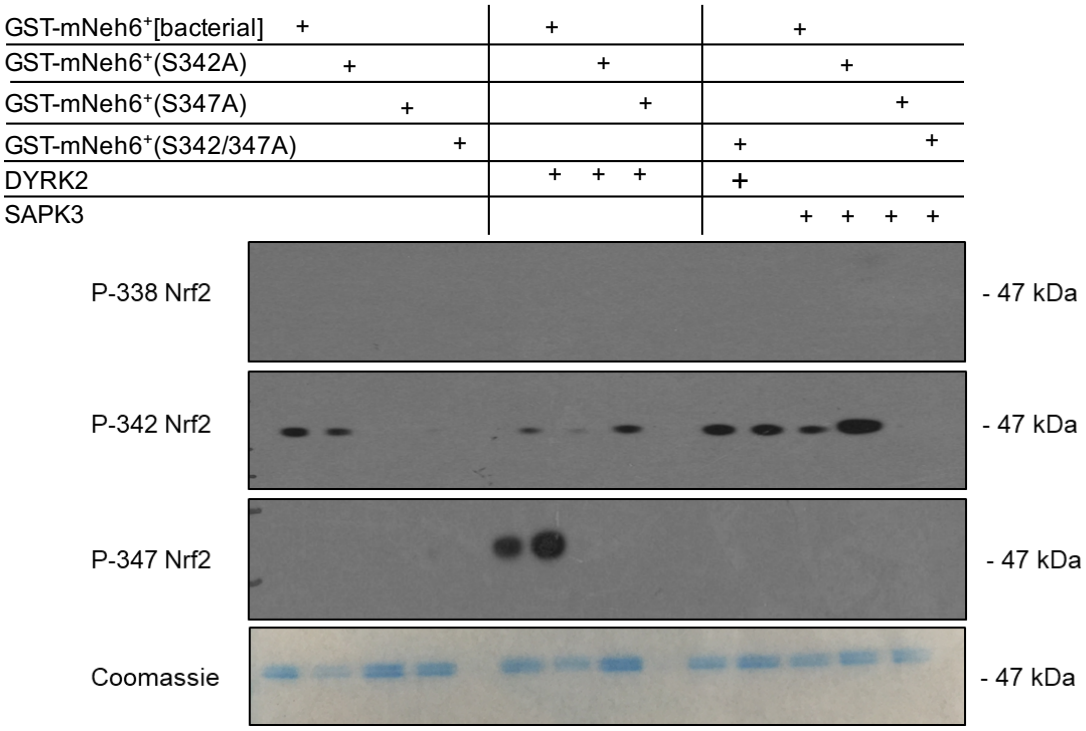


Figure 3.2.12: DYRK2 and SAPK3 phosphorylate Ser-342 and Ser-347, respectively

An immunochemical assay was carried out to determine whether DYRK2 and SAPK3 phosphorylate the same Ser residues in the Neh6 domain of mNRF2. Purified recombinant GST-tagged mNeh6^{+} protein (0.1 µg/µl), or mNeh6^{+} protein bearing a single or double point mutations (0.1 µg/µl) were generated in E.coli (BL21(DE3)pLysS) cells was incubated in a reaction mixture of 40 µl with purified activated recombinant DYRK2 or SAPK3 (each at a concentration of 2 mU/µl) for 1 hr in the presence of 10 mM $MgCl_2$ and 0.1 mM cold ATP at 30°C. The reactions were terminated by addition of SDS gel loading buffer, the samples (20 µl) were loaded on a SDS-PAGE gel, and excess ATP removed by electrophoresis. Gels were transferred onto PVDF membrane, blocked in 5% (w/v) milk- TBST and then incubated overnight with either phospho-Ser-338 (P-338), phospho-Ser-342 (P-342) or phospho-Ser-347 (P-347) NRF2 antibodies.

3.2.11 Phosphorylation of mNRF2 at Ser-338 and Ser-342 by GSK-3, and at Ser-347 by DYRK isoenzymes leads to a reduction in its transcriptional activity

To study the impact that GSK-3 and DYRK enzymes have on the transcriptional activity of mNRF2 an ARE-luciferase reporter assay was performed using pharmacological agents to inhibit kinases of interest; 5 μM MK2206 was employed to inhibit AKT, 5 μM CT99021 was employed to inhibit GSK-3, and 10 μM Harmine was employed to inhibit DYRK isoenzymes. The effects of these inhibitors on the activity of ectopic mNRF2$^{\Delta 17\text{-}32}$-V5 and mutants derived from it (i.e. S335A, S338A, S342A and S347A) was examined in human lung NRF2 KO A549 cells, which harbour mutant *KEAP1* and have had hNRF2 knocked out by CRISPR/Cas9 gene editing. As shown in Figure 3.2.13, treatment with either CT99021 or Harmine led to an increase in ARE-driven transcription by both ectopic mNRF2$^{\Delta 17\text{-}32}$-V5 and mNRF2$^{\Delta 17\text{-}32}$(S335A)-V5. These results were anticipated because it was envisaged that inhibition of either GSK-3 or a priming kinase would result in a reduction in SCF$^{\beta\text{-}TrCP}$-mediated degradation of mNRF2, leading to an increase in mNRF2 protein and thus enhanced transactivation activity. The fact that ARE-driven luciferase activity increased in NRF2 KO A549 cells transfected with mNRF2$^{\Delta 17\text{-}32}$(S335A)-V5 when treated with CT99021 or Harmine suggests that this site may not be phosphorylated by GSK-3, even though it is part of the motif to which β-TrCP binds. In the presence of MK2206, transcriptional activity of both mNRF2$^{\Delta 17\text{-}32}$-V5 and mNRF2$^{\Delta 17\text{-}32}$(S335A)-V5 in NRF2 KO A549 cells was found to be decreased. This is because when AKT is inhibited, GSK-3 remains in its active non-phosphorylated state, and therefore is able to phosphorylate NRF2, leading to SCF$^{\beta\text{-}TrCP}$-directed degradation of the transcription factor. The transcriptional activities of mNRF2$^{\Delta 17\text{-}32}$(S338A)-V5, mNRF2$^{\Delta 17\text{-}32}$(S342A)-V5 and mNRF2$^{\Delta 17\text{-}32}$(S347A)-V5 were not altered by CT99021, Harmine or MK2206. These findings are consistent with the notion that Ser-388, Ser-342 and Ser-347 are phosphorylated by GSK-3 and DYRK isoenzymes.

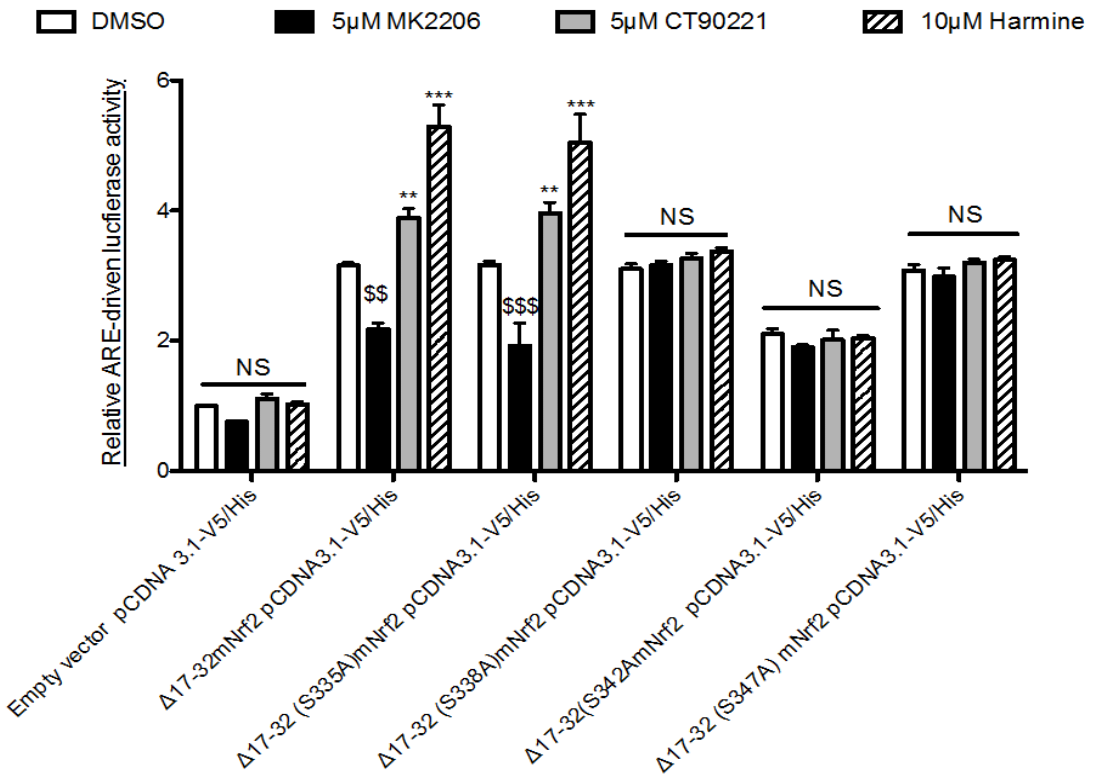


Figure 3.2.13: Phosphorylation of mNRF2 at Ser-338 and Ser-342 by GSK3 and at Ser-347 by DYRK isoenzymes diminishes transcriptional activity

*NRF2 KO A549 were co-transfected with an empty pcDNA3.1-V5/His expression plasmid or various Δ17-32 mNRF2 pcDNA3.1-V5/His mutant expression plasmids (as indicated at the bottom of the figure) along with the murine quinone reductase 1-based P-1016/nqo1-Luc ARE-containing reporter construct and the pRL-TK Renilla control plasmid. Equal amounts of DNA (2 μg) were transfected into NRF2 KO A549 cells in the different experimental groups. After overnight transfection, the cells were serum-depleted for 16 hr, followed by treatment with kinase inhibitors, or vehicle control, for a further 18 hr before the ARE-driven luciferase activity was measured. All results were normalized to renilla luciferase activity. The data are presented as follows: cells treated with DMSO in open bars; cells treated with MK2206 in black bars; cells treated with CT99021 in grey bars and cells treated with Harmine in striped bars. The data displayed are the mean values from three biological replicates and the associated SEM. T-tests were calculated using GraphPad Prism software. All significant increases are denoted with the * symbol; * represents P< 0.05, ** P<0.01 and ** P< 0.001. P values > 0.05 were deemed not significant and are denoted by "ns". Significant decreases are denoted with a $ symbol, and the same cut offs applied.*

3.2.12 Phosphorylation of mNRF2 at Ser-338 and Ser-342 by GSK-3β, and at Ser-347 by DYRK leads to a reduction in its protein stability

To determine the impact that GSK-3 and DYRK enzymes have on the stability of mNRF2 protein, a β-galactosidase assay was carried out using the mNeh6^{+}-LacZ fusion protein and its associated S335A, S338A, S342A and S347A mutants. Expression constructs for these proteins were transfected into NRF2 KO A549 cells, which were then treated with a same panel of kinase inhibitors used in Figure 3.2.13 before β-galactosidase activity was measured. As displayed in Figure 3.2.14, similar β-gal activity results were observed for the mNeh6^{+}-LacZ proteins as were obtained for the previous ARE-luciferase activity experiments. Specifically, treatment with CT99021 and Harmine increased the abundance of both mNeh6^{+}-LacZ and mNeh6^{+}(S335A)-LacZ. Whereas, treatment with MK2206 resulted in a decrease in abundance of mNeh6^{+}-LacZ and mNeh6^{+}(S335A)-LacZ, as evidenced from β-gal activity. By contrast, treatment with CT99021, Harmine or MK2206 did not affect the protein stability of either mNeh6^{+}(S338A)-LacZ, mNeh6^{+}(S342A)-LacZ or mNeh6^{+}(S347A)-LacZ.

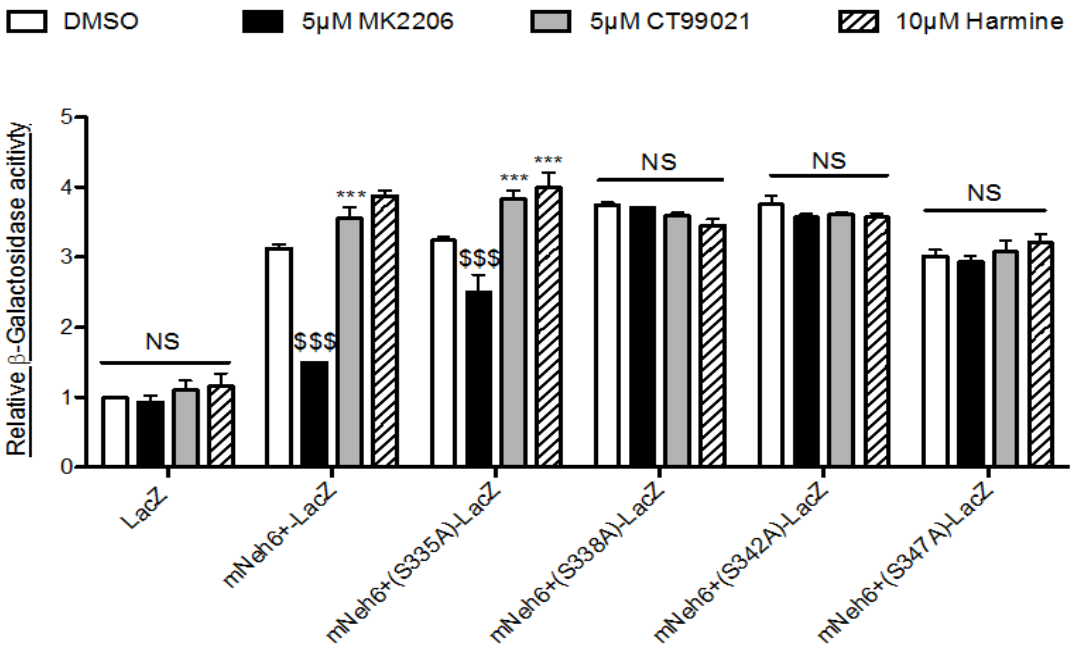


Figure 3.2.14: Phosphorylation of mNRF2 at Ser-338 and Ser-342 by GSK3 and at Ser-347 by DYRK2 leads to reduction in protein stability

*NRF2 KO A549 cells were transfected with a pcDNA3.1 expression vector for a V5-tagged fusion protein comprising mNeh6^{+} coupled at its C-terminus to LacZ (i.e. mNeh6^{+}LacZ -V5), or expression vectors for mNeh6^{+}LacZ-V5 bearing individual amino acid mutations as indicated in the figure above. As a control NRF2 KO A549 cells were transfected with the empty LacZ-V5 expression vector. After overnight transfection, the cells were serum-starved for 16 hr followed by treatment with kinase inhibitors for a further 18 hr before the β-gal activity was measured, which was later normalized against protein concentration. The data are presented as follows: cells treated with DMSO in open bars; cells treated with MK2206 in black bars; cells treated with CT99021 in gray bars and cells treated with Harmine in striped bars. The data displayed are the mean value from three biological replicates and the associated SEM. T-tests were calculated using GraphPad Prism software. All significant increases are denoted with the * symbol; * represents P< 0.05, ** P<0.01 and ** P< 0.001. P values > 0.05 were deemed not significant and are denoted by "ns". Significant decreases are denoted with a $ symbol, and the same cut offs applied.*

3.2.13 Altering GSK-3 and DYRK phosphorylation with small molecule inhibitors influences the chemosensitivity of lung cancer cells lines towards cisplatin.

To determine if regulation of NRF2 by GSK-3 or DYRK could impact cell viability, sensitivity to the platinum-based chemotherapeutic agent cisplatin was accessed using the MTT assay. Human A549 cells, which express high levels of hNRF2 due to a mutation in *KEAP1* and a hypermethylation of the *KEAP1* promoter, were treated with serial dilutions of cisplatin in the absence or presence of 5 μM MK2206, 5 μM CT99021, 1 μM Harmine or 1 μM ID-8; the latter compound is a pan DYRK family inhibitor that was included to support any effects observed with Harmine. A shown in Figure 3.2.15 and Table 3.2.2, treatment with MK2206 re-sensitizes cells to cisplatin, this is as expected because AKT inhibition will stimulate $SCF^{-\beta\text{-}TrCP}$ mediated degradation of hNRF2. Conversely, treatment with CT99021, Harmine or ID-8 enhances the resistance of A549 cells to cisplatin. However, since treatment with CT99021, Harmine and ID-8 increases the IC50 towards cisplatin making the cell more resistant, this means that the curve generated does not reach 50% percentage cell viability, even at the highest concentration of cisplatin. This means that the IC50 values given for treatments with these three inhibitors is an estimated value extrapolated from the data displayed. It should also be noted that all of the inhibitors used will have NRF2-independent affects also.

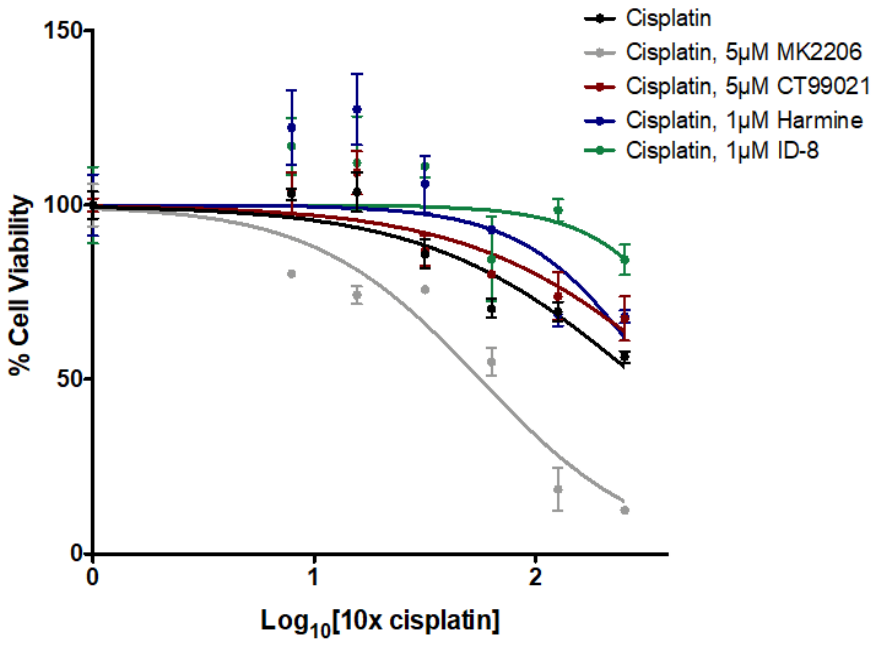


Figure 3.2.15: Sensitivity of the human lung cancer cell line A549 toward cisplatin is alerted in the presence of inhibitors that either alter GSK-3 signalling.

To determine whether sensitivity towards cisplatin might be influenced by hNRF2 through manipulation of its DSGIS-containing phosphodegron, AKT, GSK-3 and DYRK activities in A549 cells were inhibited using MK2206, CT99021 and Harmine (or ID-8), respectively, in the presence of various concentrations of cisplatin, before a MTT assay was performed. The mean percentage cell viability (from three separate wells) and the associated percentage error are displayed. These data shown are from one experiment that is representative of three independent biological replicates.

	Treatment				
Average IC50 (μM)	Cisplatin	Cisplatin, 5μM MK2206	Cisplatin, 5μM CT99021	Cisplatin, 1μM Harmine	Cisplatin, 1μM ID-8
	26.1	8.1	40.8	45.4	63.5
Stdev	3.0	2.7	7.4	5.4	6.3
Statistical significance (compared to cisplatin alone)	-	*	ns	*	**

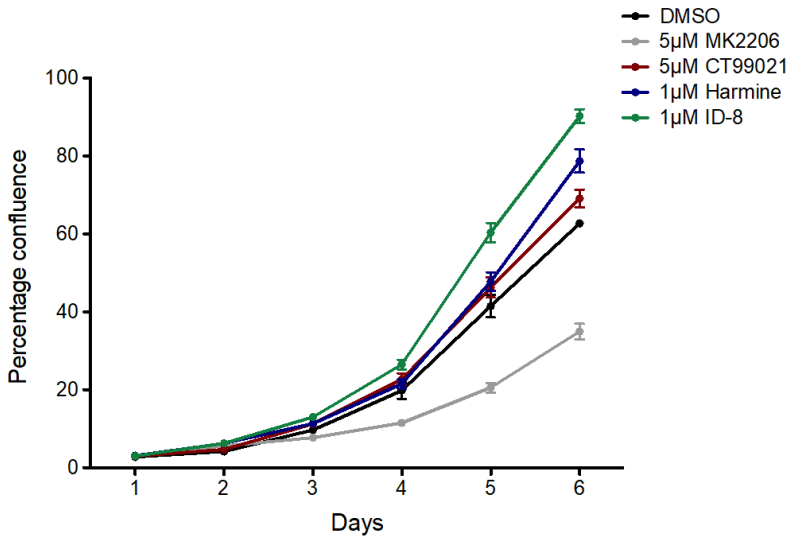
Table 3.2.2: Manipulation of GSK-3 priming kinases ability to phosphorylate hNRF2 affects chemosensitivity.

*Table showing that combining cisplatin treatment with either AKT inhibition, GSK-3 inhibition or DYRK inhibition influences sensitivity. Data displayed is the mean value from three independent biological replicates and the associated StDev. All IC50 values andT-tests were calculated using GraphPad Prism software. All significant increases (relative to the cisplatin alone data) are denoted with the * symbol; * represents P< 0.05, ** P<0.01 and ** P< 0.001. P values > 0.05 were deemed not significant and are denoted by "ns". Significant decreases are denoted with a $ symbol, and the same cut offs applied.*

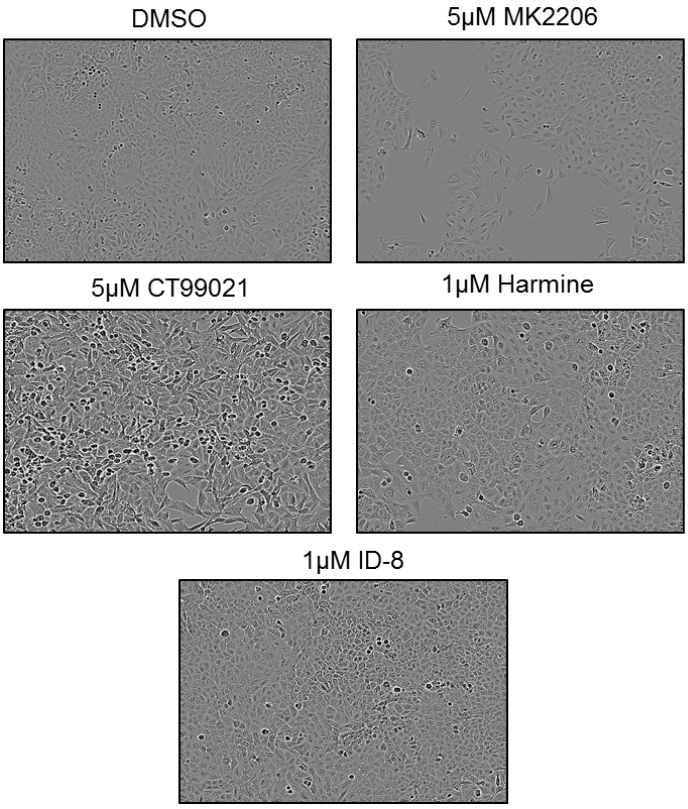
3.2.14 Alterations in GSK-3 and DYRK phosphorylation of hNRF2 affects cell proliferation

The proliferative ability of A549 cells in the presence of 5 μM MK2206, 5 μM CT99021, 1 μM Harmine or 1 μM ID-8 was accessed using the IncuCyte Zoom™ imaging system and through colony formation assays, to determine whether priming phosphorylation by DYRK or phosphodegron formation by GSK-3 altered cell proliferation. The data displayed in Figure 3.2.16, show that inhibition of AKT leads to a reduction in cell growth. Whereas, inhibition of GSK-3, or DYRK leads to an enhancement of cell proliferation. This suggests that altering the ability of GSK-3 to phosphorylate of hNRF2 through either targeting the kinase itself or targeting its associated priming kinase can affect cancer cell proliferation.

A



B


DMSO
5μM MK2206
5μM CT99021
1μM Harmine
1μM ID-8

C

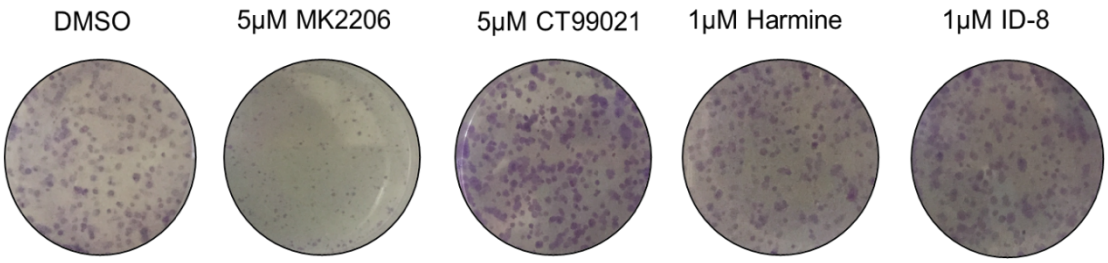

DMSO
5μM MK2206
5μM CT99021
1μM Harmine
1μM ID-8

Figure 3.2.16: Phosphorylation of hNRF2 at Ser-338 and Ser-342 by GSK3 and at Ser-347 by DYRK2 leads to enhanced proliferation.

(A) Proliferative ability was accessed using IncuCyte ZOOM™, which measures free space in a cell culture dish that can then be correlated with cell proliferation. Approximately 24 hr after seeding, A549 cells were drugged with either: DMSO, 5 μM MK2206, 5 μM CT99021, 1 μM Harmine or 1 uM ID-8. Plates were then scanned every 24 hr for 6 consecutive days and maintained in a 37°C 5% CO_2 incubator. Graphs were then generated using IncuCyte ZOOM™ software analysis package. The graph displayed is the mean value from 6 wells in one experiment and the associated SEM. This is representative of three independent biological replicates. (B) Representative images taken using the Incucyte Zoom taken from one well with each of the drug treatments on the 6th day. (C) The ability of A549 cells to form colonies was accessed through crystal violet staining. Cells were seeded at a low density, then drugged with the respective inhibitors for two weeks. After which the colonies were fixed and the plates stained. Images shown are from one experiment and are representative of three independent biological replicates.

3.3 Discussion

In this chapter the role for a priming kinase in the degradation of NRF2 through the $SCF^{\beta\text{-}TrCP}$ complex has been characterised. Phosphorylation by a priming kinase will enhanced subsequent GSK-3β mediated phosphorylation of the Neh6 domain resulting in the proteaosomal degradation of NRF2.

3.3.1 The proposed mechanism by which NRF2 undergoes a priming event before GSK-3 phosphorylation is similar to that seen with the classic β-TrCP substrate β-catenin

The degradation of β-catenin is a classic example of the required amino acid spacing between a priming kinase site and subsequent GSK-3 phosphorylation sites. As shown in Appendix 8.7, phosphorylation of Ser-45 of β-catenin by CK1 primes the substrate for successive phosphorylation by GSK-3 at residues 41,37 and 33 (Marin et al., 2003). These phosphorylation events lead to the recruitment and binding of $SCF^{\beta\text{-}TrCP}$ resulting in the ubiquitination of β-catenin and its degradation. We have demonstrated in this chapter the priming of NRF2 for phosphorylation by GSK-3 using the DYRK isoenzymes as examples. DYRK enzymes phosphorylate NRF2 at Ser-347, priming for subsequent

phosphorylation at Ser-342 and Ser-338 by GSK-3. Leading to degradation of NRF2 in the same manner as that of β-catenin.

3.3.2 Impact that commonly occurring cancer mutations will have on this pathway

Even though mutations in the genes that encode β-TrCP and GSK-3 are very rare in cancer (and appear to be redundant in their effect), KEAP1-independent degradation of hNRF2 will be altered in several tumours due to mutations in genes that influence GSK-3 activity. Mutations in *PI3KCA* and *PTEN* both lead to hyper-activation of AKT and are very commonly found in lung cancer tumours (Campbell et al., 2016). Aberrant AKT signalling will lead to increased phosphorylation of GSK-3 at Ser-9 and Ser-21, meaning that it will be inactive. Under such conditions manipulating the priming of hNRF2 will not be able to stimulate degradation. Suggesting that the mutational landscape of a tumour would have to be characterized before treatment with compounds that alter the priming of hNRF2.

3.3.3 Inhibition of DYRK will affect the priming of other substrates not just NRF2

Although targeting the priming kinase is a more selective way of reducing GSK-3 signalling than inhibiting GSK-3 itself, increasing DYRK activity will not only affect NRF2. Woods and colleagues have demonstrated the DYRK family members prime both eIF2B and tau by phosphorylating Ser-539 and Thr-212, respectively (Woods et al., 2001). eIF2B is fundamental to translation initiation in all mammalian organisms that is often elevated in cancer, providing a pro-survival advantage to the cancer cell (de la Parra et al., 2018). Tau the microtubule associated protein is often hyperphosphorylated by GSK-3 in Alzheimer's disease. This suggests that treatment with DYRK activators would be beneficial in cancer to reduce the levels of NRF2 and inactivate eIF2B and could potentially decrease tau phosphorylation also.

3.3.4 Conflicting results: DYRK1A has been shown previously to enhance the expression of the NRF2-target gene NQO1

The model we proposed in this chapter implies that an increase in DYRK activity will stimulate mNRF2 degradation and thus lead to a reduction mNRF2-target gene expression. *NQO1* is a prototypical NRF2-target gene that has been shown to be decreased upon NRF2 silencing. Work by Noll et al suggests that

DYRK1A increases nuclear mNRF2, independently of KEAP1, through either directly phosphorylating mNRF2 itself or through an indirect mechanism involving ERK1/2 (Noll et al., 2012). They make these conclusions from a series of experiments using the inhibitor Harmine, which they refer to as a DYRK1A inhibitor, however work by Bain et al. (2007) would suggest that Harmine can inhibit all DYRK isoenzymes. Also, there appears to be some variability in the expression of mNRF2 target genes which are induced upon the nuclear localization of mNRF2, with HMOX1 decreasing and NQO1 remaining unchanged in the liver of their overexpression DYRK1A mouse. Noll and colleagues attribute this to the fact the *NQO1* is not only controlled by NRF2 but also by arylhydrocarbon receptor but they overlook the fact *NQO1* can also be transcriptionally activated by c-Jun and NRF1 (Jaiswal, 2000). The decrease in HMOX-1 expression they account for by an increase in NF-κB signalling. Due to the variability in the data and that they haven't fully characterized the phosphorylation site for DYRK1A at either Ser-400 or Ser-569, at this time it is difficult to compare this data to the findings in this chapter.

3.3.5 Phosphorylation of Ser-335 is mediated by an unknown kinase

Rada and co-workers were the first group to identify the two sites to which β-TrCP binds NRF2, Ser-335 and 338 (Rada et al., 2012). They demonstrated through the use of NMR and docking experiments that GSK-3 mediates the phosphorylation of both of these sites in the Neh6 domain of mNRF2 leading to the formation of the phosphodegron to which β-TrCP binds. The data in this chapter supports the notion that GSK-3 phosphorylates Ser-338 but indicates the phosphorylation at Ser-335 is not carried out by GSK-3. Data displayed in Figure 3.2.13 and 3.2.14 show that when Ser-335 is mutated to Ala, treatment with the GSK-3 inhibitor, CT99021, still induces mNRF2 transcriptional activity and protein stability. Also, treatment the AKT inhibitor, MK2206, which will lead to an increase in GSK-3 activity, decreases both mNRF2 transcriptional activity and protein stability. This suggests that Ser-355 is not phosphorylated by GSK-3. This is also supported by the idea that the spacing between subsequent GSK-3 phosphorylation sites in a substrate is normally three amino acids, this suggests that GSK-3 would not be able to phosphorylate both Ser-335 and Ser-338 because of their only being a two amino acid gap between both sites (Dajani et al., 2001). A potential reason for the variation between our findings and the Rada

et al., paper, is that we have used single site mutant constructs and purified proteins. Whereas, they have used a double mutant construct, were both Ser-335 and 338 have been mutated to Ala. We believe that with this double mutant construct the majority of its effect is due to the mutation of the 338 site and not the 335 site. At this point we don't have any data to indicate what kinase may phosphorylate Ser-335 but the amino acid sequence around the site conforms to a classical CK1 consensus sequence, D/EXXS/T, where X is any amino acid (Marin et al., 2003).

3.3.6 Development of DYRK activators to reduce NRF2 signalling in lung cancer

The data shown here with the two pan-DYRK inhibitors, Harmine and ID-8, demonstrates that DYRK can phosphorylate the Neh6 region at Ser-347 priming for subsequent GSK-3β phosphorylation. However, inhibition of DYRK would not be advantageous for the treatment of lung cancer patients with high levels of hNRF2. As inhibition of DYRK would reduce GSK-3β phosphorylation and halt $SCF^{\beta\text{-}TrCP}$ mediated degradation of the transcription factor. Unfortunately, there is currently no activators of DYRK available which has in part impeded to progress of this research. A potential route moving forward would be to carry out DYRK isoform specific gene knock downs using siRNA but one caveat would be that when one DYRK isoform is depleted it is likely that another isoform will step into its role. Another approach would be to carry out over expression experiments, to show that over expression of DYRK enhances GSK-3β phosphorylation and leads to a reduction in hNRF2 signalling. Though care would have to be taken to ensure that overexpression did not generate an artificial scenario where DYRK levels were considerably higher than that found in cells.

3.3.7 Regulating the substrate specify of DYRK isoforms

3.3.7.1 DYRKs show isoform specific preferences for the position of an arginine residue in the substrate protein.

Work by Campbell and Proud revealed that the identity of the amino acids neighbouring a DYRK phosphorylation site dictates which member of the DYRK family that will target that substrate (Campbell and Proud, 2002). They show that both DYRK2 and DYRK3 have a preference for an Arg residue at either -2 or -3 from the DYRK phosphorylation site. However, DYRK1A will only phosphorylate when there is an Arg at position -3. The specificity for DYRK4 was not analysed

in this publication. Interestingly, when the sequence surrounding the Ser-347 phosphorylation site in the Neh6 is compared between human and mNRF2 the only difference between the two sequences is the presence of a Arg residue at -2 in mNRF2 and a Val residue at the same position hNRF2. This potentially accounts for the five-fold higher phosphorylation of the mouse peptide in comparison to the human peptide by both DYRK2 and DYRK3 in the *in vitro* kinase screen (Table 3.2.1). This highlights the importance of looking at the neighbouring amino acid sequence that surrounds a phosphorylation site to aid in the determination of the kinase which targets that site.

3.3.7.2 Localisation of DYRK isoforms

Another way by which substrate specificity is determined for DYRK isoenzymes is by subcellular localization. The localization of endogenous DYRK1A shows tissue specific variances of either cytoplasmic or nuclear. Whereas, DYRK3 is almost completely nuclear. Little is known about DYRK4 but it is thought to be both cytoplasmic and nuclear. Interestingly the subcellular localization of DYRK2 appears to be in part influenced by stress. Under genotoxic stress conditions DYRK2 translocates from the cytosol to the nucleus (Sergi et al., 2011). Since KEAP1 is predominantly cytoplasmic, it has been suggested that regulation through the Neh6 domain via β-TrCP is a nuclear event (Chowdhry et al., 2013). This is further supported by the fact that several components of the SCF degradation machinery are found in the nucleus (Lassot et al., 2001). If we assume that β-TrCP mediated degradation of NRF2 occurs in the nucleus, then we look to see which DYRK family members are nuclear. It suggests that phosphorylation of Ser-347 could be mediated by either DYRK3 or DYRK4, or potentially DYRK2 under stressed conditions. However, it is not currently known whether phosphorylation of the Neh6 region by either a priming kinase or GSK-3 occurs before translocation into the nucleus or within that subcellular compartment.

3.3.8 Concluding remarks

This work in this chapter has demonstrated that several kinases are able to phosphorylate the Neh6 region of NRF2, functioning as a priming event for subsequent GSK-3 mediated phosphorylation. DYRK has been shown to prime for GSK-3 phosphorylation of various substrates, and here we have used this family of kinases to demonstrate priming of NRF2. This provides a completely

novel way of stimulating NRF2 degradation, deceasing cell proliferation and re-sensitizing cancer cells to chemotherapeutics agents, which could be applied to the treatment of tumours with mutations in *KEAP1* or *NFE2L2.* However, the work done here does add to the ongoing argument that GSK-3 inhibition has the potential to be oncogenic. Therefore, it would be advantageous to produce either small molecules that target DYRK isoforms increasing their expression or the primed phosphate to which GSK-3 binds, mimicking a constitutively primed GSK-3.

4. Bioinformatics analysis of the prevalence, co-occurrence and impact of mutations in *KEAP1* and *NFE2L2* in lung cancer cell lines and tumours.

4.1 Introduction

Cancer does not develop as a consequence of a single mutation but rather involves several gain-of-function mutations in oncogenic genes and loss-of-function mutations in tumour suppressor genes occurring in the one cell (Hanahan and Weinberg, 2000; Hanahan and Weinberg, 2011).

Historically, lung cancer was an extremely rare disease pre-twentieth century but has increased dramatically in prevalence in recent decades and is now the leading cause of cancer associated mortality worldwide with an estimated 1 million deaths per year (Evans, 2013; Torre et al., 2015; The Cancer Genome Atlas Research, 2014). Cancer research UK (CRUK) has named lung cancer as one of the four cancers of substantial unmet need due to its associated poor 5-year survival rate and modest advancements in diagnosis and treatment in the past decade.

4.1.1 Treatment of lung cancer is dependent on stage of diagnosis

The effectiveness and treatment type available to combat lung cancer is crucially dependant on the stage of the cancer at time of diagnosis (Shiono et al., 2015). Early stage cancer can be diagnosed through either computed tomography (CT) or bronchoscopy and treatment of the disease is often very effective (Nanavaty et al., 2014). There are four main stages of lung cancer numbered I-IV. Stage I lung cancer has the highest overall survival rate as patients may be treated by surgical resection which offers the lowest rate of recurrence. Stage II cancer can also be surgically removed in appropriate patients but adjuvant chemotherapy is given post-operatively to ensure complete removal of the disease (Dasari and Tchounwou, 2014; Hensing, 2014). Adjuvant chemotherapy for lung cancer often takes the form of a platinum-doublet regime in which the platinum-based compound is often cisplatin, in combination with a third-generation chemotherapeutic agent, such as docetaxel, gemcitabine or paclitaxel (Baggstrom et al., 2007; Horita et al., 2017). Cisplatin is a cytotoxic

anti-neoplastic compound that interacts with DNA, causing DNA cross-links and the formation of adducts, which halt DNA replication and leads to the activation of several signalling cascades that induce apoptotic cell death (Dasari and Tchounwou, 2014; Siddik, 2002). Chemotherapy targets highly replicative cells, such as cancer cells which have increased metabolic demand, but will also affect non-cancerous cells. This off-target effect may account for some of the side effects associated with chemotherapy in patients, such as vomiting and hair loss, as gastrointestinal and hair follicle cells replicate extensively (Evans, 2013).

However, the majority of lung cancer diagnosed is at a late stage (stage IV). Presentation is delayed for a number of reasons; the disease is largely asymptomatic in the early stages, symptoms when present are poorly defined and often attributed to other diseases, and a significant proportion of patients fear a cancer diagnosis and may mask initial symptoms (Evans, 2013). An estimated 80% of patients present with advanced stage disease that is no longer curative (Tod et al., 2008). At this later stage, surgical intervention is not possible and often palliative chemotherapy is the only option.

4.1.2 Histological subtypes of lung cancer

Lung cancer can be broadly divided into two main histological groups: small cell lung cancer (SCLC) which accounts for 15% of all lung cancer cases and non-small cell lung cancer (NSCLC) which accounts for the remaining 85% (Evans, 2013). SCLC has a stronger disease focal point of origin arising in the larger airways and presents as a very aggressive rapidly proliferative tumour with extensive metastasis to distal sites (Nanavaty et al., 2014). NSCLC is often found in the peripheral tissue and can be histologically further subdivided, according to the tissue of origin, into adenocarcinoma (which is the most prevalent), squamous carcinoma and large cell carcinoma (Pikor et al., 2013).

4.1.3 The role that environmental factors play in the development of lung cancer

Cigarette smoke has been linked to the development of lung cancer and poor disease outcome for decades (Shiono et al., 2015), with smokers being 10-20 times more likely to develop the disease than non-smokers (Hensing, 2014). The popularity associated with smoking is a relatively new cultural phenomenon and maybe in part explains why this disease was very rare before the twentieth century (Torre et al., 2015). The carcinogenic effects of smoking accumulate over

time and hence lung cancer is more prevalent in the elderly population. Smoking is not the sole risk factor associated with lung cancer, as exposure to other environmental factors also contributes to the development of the disease, including: asbestos, radon gas, arsenic and silica (Hensing, 2014; Evans, 2013). These additional environmental factors, along with exposure to second-hand cigarette smoke, may play a role in the development of lung cancer in non-smokers. Lung cancer, predominantly of the NSCLC histological type, has been diagnosed in “never-smokers” and in these cases, is thought to be driven by genetic alterations and the types of environmental factors listed above. Never-smoker is a term given to an individual who has consumed under 100 cigarettes in their lifetime (Torre et al., 2015).

4.1.4 The role that genetic factors play in the development of lung cancer

Research into the role played by genetic alterations in development of NSCLC has rapidly progressed due to the high frequency (approximately 10-15 % of cases) of this lung cancer subtype occurring in never-smokers (Samet et al., 2009). However, the identification of novel driver genes in NSCLC has been extremely challenging due to its associated high somatic mutation rate that leads to the accumulation of passenger mutations. Lung adenocarcinoma (ADC) and lung squamous cell carcinoma (SQCC), respectively the first and second most common forms of NSCLC, have both shared and unique genetic landscapes (Campbell et al., 2016). Comprehensive molecular profiling of both ADC and SQCC revealed thirty-eight significantly mutated genes in ADC and twenty significantly mutated genes in SQCC, with six commonly mutated genes in both lung cancer subtypes, these include *TP53, RB1, ARID1A, CDKN2A, PIK3CA* and *NF1* (Campbell et al., 2016). Due to the few overlapping mutations, it is thought that somatic mutations in driver genes are different between ADC and SQCC. These divergences become clearer when the patterns of mutations in two lung cancer subtypes are analysed independently.

Molecular profiling by ‘The Cancer Genome Atlas’ (TCGA) consortium of 230 ADC samples from untreated patients revealed eighteen significantly mutated genes (The Cancer Genome Atlas Research, 2014), as shown in Table 4.1.1. These included, but were not limited to, *TP53* (mutated in 46%), *EGFR* (mutated in 14%), *KRAS* (mutated in 33%), *PIK3CA* (mutated in 7%), *STK11* (mutated in 17%) an *KEAP1* (mutated in 19%). Interestingly, when pathway

analysis was carried out to determine which cell signalling pathways are commonly dysregulated in ADC due to mutations, it was highlighted that alterations in oxidative stress pathways occur in 22% of tumours due to mutations in *KEAP1* (mutated in 19%), *CUL3* (mutated in less an 1%) and *NFE2L2* (mutated in 3%) (The Cancer Genome Atlas Research, 2014).

Gene name	Mutational frequency (%)
TP53	46
KRAS	33
KEAP1	17
STK11	17
EGFR	14
NF1	11
BRAF	10
SETD2	9
RMB10	8
MGA	8
MET	7
ARID1A	7
PIK3CA	7
SMARCA4	6
RB1	4
CDKN2A	4
U2AF1	3
RIT1	2

Table 4.1.1: Identification of commonly mutated genes in lung adenocarcinoma

Summary of the 18 significantly mutated genes identified by the TCGA (The Cancer Genome Atlas Research, 2014). The left-hand column of the table displays the gene name and the right-hand column is the percentage mutational frequency for each gene out of the 230 ADC samples analysed.

Analysis of the mutational landscape of 178 SQCC samples (Table 4.1.2) from untreated patients revealed ten significantly mutated genes, including *TP53* (mutated in 81%), *CDKN2A* (mutated in 15%), *PTEN* (mutated in 15%), *PIK3CA* (mutated in 16%), *KEAP1* and *NFE2L2* (The Cancer Genome Atlas Research, 2012). These data demonstrated that SQCC have much fewer mutations in *EGFR* and *KRAS*, which are very commonly mutated in ADC, but higher frequencies of mutation in *PTEN* and *PIK3CA.* However, they revealed that dysregulation in oxidative stress pathways, as seen in the ADC cases, was similarly observed in 34% of SQCC tumours, due to mutations (or copy number alterations) in *KEAP1* (mutated in 12%), *CUL3* (mutated in 7%) and *NFE2L2* (mutated in 19%). The study also highlighted that mutations in *KEAP1* and *CUL3* are often loss-of-function mutations that occur mutually exclusive to mutations in *NFE2L2* (The Cancer Genome Atlas Research, 2012).

Even though there is a large genetic divergence between ADC and SQCC, both subtypes of NSCLC show dysregulation of oxidative stress pathways due to mutations in *KEAP1, NFE2L2* and *CUL3,* three genes that are fundamental to homeostatic redox signalling.

Gene name	Mutational frequency (%)
TP53	81
CDKN2A	20
PTEN	16
PIK3CA	15
KEAP1	15
MLL2	12
HLA-A	8
NFE2L2	8
NOTCH1	7
RB1	3

Table 4.1.2: Identification of commonly mutated genes in lung squamous cell carcinoma

Summary of the ten significantly mutated genes identified by the TCGA (TCGA,2012). The left-hand column of the table displays the gene name and the right-hand column is the percentage mutational frequency for each gene out of the 178 SQCC samples

4.1.5 Impact of mutations that dysregulate the KEAP1-NRF2 axis

Due to their increased metabolic demand and the switch from oxidative phosphorylation to glycolytic metabolism, cancer cells themselves produce a large amount of ROS (Vander Heiden et al., 2009; Nogueira and Hay, 2013). To withstand this highly oxidative environment, cancer cells often activate the NRF2 signalling pathway to neutralized ROS and thus inhibit apoptosis and senescence. One way by which cancer cells enhance NRF2 activity is through somatic mutations in *NFE2L2* and *KEAP1,* which lead to increased NRF2 activity through impairing KEAP1-mediated proteasomal degradation of the transcription factor. Also the expression of oncogenic mutations in primary murine cells has been shown to directly enhance NRF2 activation and decrease levels of intracellular ROS (DeNicola et al., 2011).The exact stage during tumourigenesis at which *KEAP1* or *NFE2L2* become mutated is still unknown, but it has been suggested that it is unlikely to occur during cancer initiation, but rather during cancer progression (Hayes and McMahon, 2009).

4.1.5.1 Mutations in *NFE2L2*

Somatic mutations in *NFE2L2* have been documented in both lung cancer cell lines and in tissues. They have been shown to correlate with poor prognosis and are more predominantly found in SQCC than ADC (Sasaki et al., 2010; Sasaki et al., 2013a). Mutations in *NFE2L2* are thought to be gain-of-function mutations resulting increased copy number expression of the transcription factor (Sasaki et al., 2012). Somatic mutations in *NFE2L2* are clustered, with a clear focal enrichment point. This type of ‘hot spot’ mutational pattern is characteristic of mutations in oncogenes, which typically occur within a small region of the gene (Baeissa et al., 2016). Mutations in these regions will have a pronounced effect on NRF2 signalling as NRF2 does not have to undergo posttranslational modification for KEAP1 to bind to it, and so the structural integrity of these sites is fundamental for homeostatic NRF2 regulation Therefore, mutation of the DLG and ETGE motifs will lead to constitutive activation of NRF2 independently of oxidative stress (Shibata et al., 2008).

Work by TCGA revealed that mutations in *KEAP1, NFE2L2* and *PTEN* never co-occur in the same tumour (The Cancer Genome Atlas Research, 2012). Mutations in *NFE2L2* have been demonstrated to co-occur with *PIK3CA* mutations (Scheffler et al., 2014) but an extensive analysis of the co-occurrence/

mutual exclusivity of *NFE2L2* mutations with mutations in other lung cancer related genes is lacking.

4.1.5.2 Mutations in *KEAP1*

Unlike mutations in *NFE2L2*, somatic mutations in *KEAP1* are found spread throughout the protein, which represents the mutational patterned associated with tumour suppressor genes, and are often associated with low *KEAP1* expression (Hast et al., 2014; Ohta et al., 2008). Due to their wide distribution, it is difficult to attribute a function to all the mutations in *KEAP1*. Another point of divergence between *KEAP1* and *NFE2L2* mutations, is that those in *KEAP1* are commonly associated with ADC (Sasaki et al., 2013b). Research has also highlighted that several mutations in *KEAP1* are G>T transversions, which are characteristic mutations associated with exposure to tobacco smoke. This may in part explain why *KEAP1* mutations often co-occur with *KRAS* mutations in ADC, which are also correlated with high smoking incidence (Singh et al., 2006; Ahrendt et al., 2001; Romero et al., 2017).

Work by Ben Major and colleagues, demonstrated that the majority of mutations in *KEAP1* do not affect the binding and ubiquitination of NRF2 but somehow halt proteasomal degradation of the transcription factor (Hast et al., 2014). Passenger mutations were also identified in *KEAP1* that do not appear to influence NRF2 degradation. Also, a small population of *KEAP1* mutations were actually shown to decrease NRF2 activity through increased binding of KEAP1 to NRF2, these mutations where given the name “super binders” or additionally NRF2-complexed hypomorph (ANCHOR) mutations. It is thought that the presence of theses mutations leads to a stabilisation of KEAP1 and a specific increase in NRF2 expression but the function of such mutations is still not fully understood (Hast et al., 2014; Cloer et al., 2018). One of the biological functions of KEAP1 is to serve as a redox sensor due to the presence of reactive cysteine residues such as, Cys-151, Cys-273 and Cys-288, which surprisingly are not found to be mutated in lung cancer (Hayes and McMahon, 2009). Overall *KEAP1* mutations have been shown to lead to loss of function of KEAP1 or disruption of the molecular interaction between KEAP1 and NRF2, both of which result in enhanced nuclear accumulation of NRF2 and increased expression of NRF2-target genes (Ohta et al., 2008; Padmanabhan et al., 2006).

4.1.5.3 Mutations in *CUL3*

In comparison to mutations in *KEAP1* and *NFE2L2* that have been extensively studied for their role in lung cancer, *CUL3* mutations do not seem to have attracted much attention and are poorly understood. Results of a detailed genetic screen to identify lung cancer drivers in mice revealed that *CUL3* may function as a tumour suppressor in cooperation with mutant *PTEN* (Dorr et al., 2015). This is supported by the fact that reduced expression of CUL3 protein and mRNA has been documented in several lung cancer cell lines and tumours harbouring mutations in *CUL3.* Gene siRNA knockdown studies have indicated that a *CUL3* mutation may enhance tumour proliferation through stimulating the NRF2–signalling pathway and therefore inducing expression of antioxidant genes (Loignon et al., 2009).

Breast cancer cell lines and tumours often show decreased NRF2 expression and higher ROS levels, which can be mimicked in cells through over-expression of CUL3 (Loignon et al., 2009), suggesting that CUL3 negatively regulates NRF2 expression. Studies into papillary renal cell carcinoma (PRCC) show that mutations in *CUL3* lead to a loss of function of the gene, which elevates NRF2-signaling in PRCC tumours (Ooi et al., 2013).

4.1.6 The role that mutations in the NRF2 pathway play in enhancing cell proliferation and the development of chemo-resistance

The development of drug resistance is a huge therapeutic hurdle in the treatment of NSCLC and is one of the major reasons for the poor outcome associated with the disease, with the response rate for first line chemotherapy in patients only being 30% (Homma et al., 2009). Many mechanisms are implicated in the development of drug resistance such as: increased expression of ATP-dependant drug transporters leading to increased drug efflux out of the cell, insensitivity to apoptosis and increased expression of drug detoxification genes (Gottesman, 2002). All pathways implicated in the development of drug resistance involve genes that are transcriptionally regulated through NRF2. Indeed, siRNA-mediated knockdown of NRF2 has been shown to increase sensitivity to anti-cancer drugs and enhance cell proliferation (Homma et al., 2009).

Platinum-based chemotherapeutics function to damage DNA and stimulate ROS production leading to cancer cell death. However, resistance to

platinum chemotherapeutics can arise through several NRF2-dependent routes (Choi and Kwak, 2016). Upregulation of NRF2 has been shown to decrease levels of intracellular ROS, leading to reduced apoptosis and increased cancer cell survival. Also, *GCLC, GCLM, GPX2* and several *GSTs*, key genes involved in the biosynthesis and utilization of glutathione, are all transcriptionally regulated by NRF2. Increased levels of NRF2 lead to increased intracellular glutathione levels and increased glutathione-conjugation to drugs, including anti-cancer agents (Lan et al., 2018). The conjugation of glutathione with drugs is catalysed by GST isoenzymes, and functions to reduce toxicity and intracellular concentrations of the drug through efflux of the conjugated drugs via pumps, such as multidrug resistance-associated protein 1(MRP1) (Gamcsik et al., 2012) (The importance of glutathione-conjugation is discussed in greater detail in Chapter 5). Several ATP-dependent drug-efflux pumps, including MRP1, are regulated by NRF2 (Ji et al., 2013). The presence of mutations in *NFE2L2* and *KEAP1* leading to high levels of NRF2 protein and increased expression of MRP transporters (Mahaffey et al., 2012). Expression of the NRF2-target gene *AKR1C1* has also been linked with enhanced drug resistance. With siRNA mediated knockdown of *AKR1C1* in the *KEAP1* mutant NSCLC cell line, A549, leading to a marked reduction in resistance to cisplatin (Chen et al., 2017). Also *HMOX-1*, a NRF2-target gene associated with stress responses has been shown to be increased in pancreatic tumours, and depletion of *HMOX-1* gene expression leads to re-sensitization of these cells (Berberat et al., 2005).

Rapid cellular proliferation is one of the major hallmarks of cancer and is often due to dysregulation of normal cell cycle events (Feitelson et al., 2015). High levels of NRF2, due to mutations/hypermethylation of *KEAP1*, has been linked to enhanced cell proliferation. This is thought to be due to induction of a G1 phase arrest in the cell cycle (Homma et al., 2009). Additionally, several NRF2 effector genes are involved in the regulation of metabolic pathways, which in the presence of aberrant NRF2 may lead to enhanced proliferation (Mitsuishi et al., 2012).

The role that mutations in *KEAP1* and *NFE2L2,* and to a later extent *CUL3*, play in the development of lung cancer has been extensively investigated. Also, the impact that NRF2 plays in the development of drug resistance through its diverse range of downstream target genes has been very well characterized.

However, an extensive analysis into the co-occurrence of *KEAP1/NFE2L2/ CUL3* mutations with other mutations, the zygosity of these mutations and the effect that these mutations have on the expression of NRF2-target genes associated with the development of resistance is lacking. In this chapter, we have employed a bioinformatics approach to address these points and demonstrate dramatic differences in the context and consequences of mutations in *KEAP1*, *NFE2L2* and *CUL3*.

4.2 Results

4.2.1 *KEAP1* and *KRAS* are the two most commonly mutated genes in lung cancer cell lines, whereas tumours harbour a high number of *PIK3CA* and *KRAS* mutations

Initially we investigated the occurrence of mutations in genes involved the NRF2-signalling pathway in both lung cancer cell lines and tumours. Using the software TICBO Spotfire®, we combined and filtered the Sanger, Cancer cell line encyclopaedia (CCLE) and AstraZeneca datasets for lung cancer cell lines. At this point we did not further subdivide the datasets to analyse specific histological subtypes of lung cancer but looked at the disease as a whole. This gave us 239 lung cancer cell lines in total, which we then filtered for those harbouring mutations in *KEAP1, NFE2L2, CUL3, KRAS, EGFR, PTEN, PIK3CA* and *STK11,* and then calculated the percentage of cell lines containing each mutation. We chose to study the impact of these mutations because the literature suggests that they have the potential to alter NRF2 degradation.

Interrogation of these datasets revealed that amongst the eight genes examined, mutations in *KRAS* (20%) and *KEAP1* (12%) were the top most common events amongst the 239 human cell lines (Table 4.2.1A). Moreover, interrogation of the human pan-lung cancer tumour dataset, which included 1144 tumours of mixed histology (Campbell et al., 2016), showed mutations in *PIK3CA* (24%) and *KRAS* (23%) were the two most common events (Table 4.2.1B). The increased frequency of *PIK3CA* mutations in the tumour data set in comparison to the cell line database may be due to the tumour dataset containing a larger proportion of SQCC samples, which often have mutations in this gene.

Together, these data suggest that in lung cancer, mutations in *KEAP1* arise more commonly than mutations in either *NFE2L2* or *CUL3,* and that mutations in *NFE2L2* are more common than mutations in *CUL3*.

A

Cell lines			
Gene	Frequency of alteration		Ranking according to prevalence
	In alerted group	In unaltered group	
KEAP1	28 (12%)	211 (88%)	2
NFE2L2	8 (3%)	231 (97%)	5
CUL3	8 (3%)	231 (97%)	6
KRAS	47 (20%)	192 (80%)	1
EGFR	15 (6%)	224 (94%)	4
PTEN	15 (6%)	224 (94%)	4
PIK3CA	16 (7%)	223 (93%)	3
STK11	15 (6%)	224 (94%)	4

B

Tumour samples			
Gene	Frequency of alteration		Ranking according to prevalence
	In alerted group	In unaltered group	
KEAP1	170 (15%)	974 (85%)	3
NFE2L2	105 (9%)	1039 (91%)	6
CUL3	51 (4%)	1093 (96%)	7
KRAS	259 (23%)	885 (77%)	2
EGFR	164 (14%)	980 (86%)	4
PTEN	102 (9%)	1042 (92%)	6
PIK3CA	276 (24%)	868 (76%)	1
STK11	1188 (10%)	1026 (90%)	5

Table 4.2.1: Frequency of mutations in genes associated with the NRF2-signalling pathway in lung cancer

(A) Incidence of mutations in KEAP1, NFE2L2, CUL3, KRAS, EGFR, PTEN, PIK3CA and STK11 in the human lung cancer cell line dataset comprising of 239 cell lines; percentages are shown in parentheses. Mutations are ranked according to their prevalence in the right-hand column of the table. (B) Incidence of mutations in KEAP1, NFE2L2, CUL3, KRAS, EGFR, PTEN, PIK3CA and STK11 in the human pan-lung cancer tumour dataset of 1144 tumours; percentages are shown in parentheses. Mutations are ranked according to prevalence in the right-hand column of the table.

4.2.2 Adenocarcinoma tumours have a higher frequency of mutations in *KRAS* and *KEAP1* than do squamous tumour samples

To determine if the frequency of mutations in genes that cause dysregulation of NRF2-signalling vary between different histological subtypes of NSCLC, we carried out the same analysis for the eight selected genes, as in section 4.2.1, but in separate ADC (The Cancer Genome Atlas Research, 2014) and SQCC (The Cancer Genome Atlas Research, 2012) datasets. This revealed a high frequency of mutation in *KRAS* (82%), *KEAP1* (19%) and *STK11* (19%) amongst 230 ADC specimens in the dataset (Table 4.2.2A). The high frequency of mutations in both *KEAP1* and *KRAS* genes is possibly not surprising because they have been reported to frequently occur in ADC (The Cancer Genome Atlas Research, 2014; Sasaki et al., 2013b); however, what is surprising is the low frequency of *EGFR* mutations, which have been shown to be extremely prevalent in ADC (The Cancer Genome Atlas Research, 2014). By contrast with ADC, amongst the 178 SQCC specimens, mutations in *PIK3CA* (48%), *NFE2L2* (19%) and *KEAP1* (14%) were commonly observed (Table 4.2.2B). This correlates with the data produced by TCGA showing the *PIK3CA* mutations are more prevalent in SQCC than ADC and often occur alongside *NFE2L2* mutations (The Cancer Genome Atlas Research, 2012). Also, it supports the notion that mutations in *KEAP1* are more common in ADC than mutations in *NFE2L2* but in SQCC the opposite is observed, with *NFE2L2* mutations being more common than those in *KEAP1*. *CUL3* mutations rarely occur in either of the histological subtypes of NSCLC. It was also noted that the frequency of *PTEN* mutations was low in both tumour data sets, and *CUL3* mutations are thought to occur alongside *PTEN* mutations.

A

Adenocarcinoma tumour samples			
Gene	Frequency of alteration		Ranking according to prevalence
	In alerted group	In unaltered group	
KEAP1	44 (19%)	186 (81%)	2
NFE2L2	10 (4%)	220 (96%)	6
CUL3	6 (2.6%)	228 (97.4%)	8
KRAS	82 (36%)	148 (64%)	1
EGFR	40 (17%)	190 (83%)	4
PTEN	7 (3%)	223 (97%)	7
PIK3CA	20 (9%)	210 (91%)	5
STK11	43 (19%)	187 (81%)	3

B

Squamous tumour samples			
Gene	Frequency of alteration		Ranking according to prevalence
	In alerted group	In unaltered group	
KEAP1	25 (14%)	153 (86%)	3
NFE2L2	33 (19%)	145 (81%)	2
CUL3	14 (8%)	164 (92%)	6
KRAS	6 (3%)	175 (98%)	7
EGFR	16 (9%)	162 (91%)	5
PTEN	20 (11%)	153 (89%)	4
PIK3CA	86 (48%)	92 (52%)	1
STK11	3 (1.7%)	175 (98.35)	8

Table 4.2.2: Frequency of mutations in genes associated with the NRF2-signalling pathway in ADC and SQCC tumors

(A) Incidence of mutations in KEAP1, NFE2L2, CUL3, KRAS, EGFR, PTEN, PIK3CA and STK11 in human lung ADC tumour dataset comprising of 230 specimens; percentages are shown in parentheses. Mutations are ranked according to prevalence in the right-hand column. (B) Incidence of in KEAP1, NFE2L2, CUL3, KRAS, EGFR, PTEN, PIK3CA and STK11 in human lung SQCC tumour dataset comprising of 178 specimens; percentages are shown in parentheses. Mutations are ranked according to prevalence in the right-hand column.

4.2.3 Mutations in *KEAP1* co-occur with mutations in *KRAS* and *STK11* in both lung cancer cell lines and tumours

Identifying which mutations tend to co-occur or are mutually exclusive with mutations in *KEAP1* may give a greater understanding of the role that *KEAP1* mutations play in carcinogenesis.

To determine whether somatic mutations in *KEAP1* co-occur with mutations in *KRAS, STK11, PTEN, PIK3CA, NFE2L2, CUL3* and *EGFR*, the same lung cancer dataset of 239 cell lines mentioned previously was used. Filtering the 239 cell lines for those harbouring mutations in *KEAP1* yielded 28 cell lines, and these were then screened for additional mutations in the seven other genes of interest. This revealed *KEAP1* mutations have a strong co-occurrence with *KRAS*, *STK11* and *PIK3CA* mutations, with positive log ratio values (Table 4.2.3A). The existence of *KEAP1* and oncogenic driver mutations in *KRAS* has been documented before (Romero et al., 2017). The fact that mutations in *KEAP1* and *KRAS* occur alongside mutations in *STK11* is to be expected as *STK11* mutations have been shown to co-exist with *KRAS* mutations in ADC (The Cancer Genome Atlas Research, 2014). *KEAP1* mutations were shown to not occur alongside mutations in *PTEN, NFE2L2, CUL3* and *EGFR*, as revealed by negative log ratio values. The notion that *KEAP1* and *NFE2L2* mutations do not occur together has been documented in the literature before. However, the idea that *CUL3* mutations do not occur with *KEAP1* mutations has not been documented before, but these data are not very robust due to the low frequency of *CUL3* mutations in cell lines.

To determine whether the same patterns of co-existence of *KEAP1* mutations with other mutations are seen in both cell lines and tumours, the same analysis was carried out using a human pan-lung cancer dataset (Campbell et al., 2016). This dataset contains 170 tumours with mutations in *KEAP1*. Amongst these 170 tumour samples, *KEAP1* mutations were found to co-occur with mutations in *KRAS, STK11* and *CUL3,* and are mutually exclusive with mutations in *PTEN, PIK3CA, NFE2L2* and *EGFR* (Table 4.2.3B). These data slightly deviate from the findings for the cell line dataset, but the tumour dataset contains more data points, making it more robust.

Interestingly mutations in *KEAP1* were found to not occur alongside mutations in *EGFR* in either dataset, which is commonly mutationally activated in ADC.

A

KEAP1 mutant cell lines				
Gene	Frequency of alteration		Log ratio	Direction/Tendency
	In altered group (Total *KEAP1* mutants)	In unaltered group (Total cell lines)		
KRAS	12 (42.8%)	16 (6.7%)	2.68	Co-occurrence
STK11	5 (17.8%)	23 (9.6%)	0.89	Co-occurrence
PTEN	0 (0%)	28 (11.7%)	-	Mutual exclusivity
PIK3CA	3 (10.7%)	25 (10.5%)	0.03	Co-occurrence
NFE2L2	0 (0%)	28 (11.7%)	-	Mutual exclusivity
CUL3	1 (3.6%)	27 (11.3%)	-1.65	Mutual exclusivity
EGFR	2 (7.1%)	26 (10.9%)	-0.62	Mutual exclusivity

B

KEAP1 mutant tumour samples				
Gene	Frequency of alteration		Log ratio	Direction/Tendency
	In altered group (Total *KEAP1* mutants)	In unaltered group (Total cell lines)		
KRAS	41 (24.1%)	181 (18.6%)	0.38	Co-occurrence
STK11	39 (22.9%)	72 (7.4%)	1.63	Co-occurrence
PTEN	6 (3.5%)	61 (6.3%)	-0.83	Mutual exclusivity
PIK3CA	12 (7.1%)	82 (8.4%)	-0.25	Mutual exclusivity
NFE2L2	4 (2.3%)	80 (8.2%)	-1.80	Mutual exclusivity
CUL3	6 (3.5%)	33 (3.4%)	0.06	Co-occurrence
EGFR	4 (2.4%)	113 (11.6%)	-2.30	Mutual exclusivity

Table 4.2.3: Co-occurrence of mutations in KEAP1 with mutations in other genes that are associated with the NRF2-signalling pathway

(A) The 28 KEAP1 mutant lung cancer cell lines were filtered for the presence of mutations in KRAS, STK11, PTEN, PIK3CA, NFE2L2, CUL3 and EGFR. Mutations with a positive log ratio are designated as tending to co-occur with KEAP1 mutations, whereas those with a negative log ratio are designated as having a tendency of mutual exclusivity with mutations in KEAP1. (B) 1144 pan-lung tumour samples were filtered for the expression of KEAP1 mutations, yielding 170 KEAP1 mutant tumours. Then further sub-filtered for the presence of mutations in KRAS, STK11, PTEN, PIK3CA, NFE2L2, CUL3 and EGFR. Log ratios and tendency of mutations to occur with KEAP1 mutations were assigned as indicated above.

4.2.4 Mutations in *NFE2L2* co-occur with mutations in *PIK3CA* and *CUL3* in both lung cancer cell lines and tumours

To determine whether somatic mutations in *NFE2L2* co-occur alongside other mutations that influence the degradation of NRF2, both the lung cancer cell line and tumour datasets were filtered for those that harbour mutations in *NFE2L2*. The eight cell lines with mutant *NFE2L2* and the 105 tumours with mutant *NFE2L2* (described in Campbell et al. (2016)) were then screened for the presence of additional mutations in *KRAS, STK11, PTEN, PIK3CA, KEAP1, CUL3* and *EGFR.* This revealed that in both the cancer cell lines and in tumours mutations in *NFE2L2* co-occurred alongside mutations in *PIK3CA* and *CUL3*, with positive log ratio values in both lung cancer cell lines and tumours (Table 4.2.4A and B). The idea that *CUL3* mutations occur alongside *NFE2L2* mutations is surprising as they seem likely to result in similar effects, inhibition of NRF2 degradation through the 26S proteasome.

Interestingly, mutations in *NFE2L2* were found to be mutually exclusive from mutations in *KRAS* and *STK11* in both the cell line and tumour datasets. As noted previously, *KRAS* and *STK11* mutations co-exist with mutations in *KEAP1*. Since mutations in *NFE2L2* and *KEAP1* do not occur together, they co-occur with different oncogenic driver genes.

A

***NFE2L2* mutant cell lines**				
Gene	Frequency of alteration		Log ratio	Direction/Tendency
	In altered group (Total *NFE2L2* mutants)	In unaltered group (Total cell lines)		
KRAS	0 (0%)	8 (3.3%)	-	Mutual exclusivity
STK11	0 (0%)	8 (3.3%)	-	Mutual exclusivity
PTEN	0 (0%)	8 (3.3%)	-	Mutual exclusivity
PIK3CA	1 (13%)	7 (2.9%)	2.16	Co-occurrence
KEAP1	0 (0%)	8 (3.3%)	-	Mutual exclusivity
CUL3	1 (13%)	7 (2.9%)	2.16	Co-occurrence
EGFR	0 (0%)	8 (3.3%)	-	Mutual exclusivity

B

NFE2L2 mutant tumour samples				
Gene	Frequency of alteration		Log ratio	Direction/ Tendency
	In altered group (Total *NFE2L2* mutants)	In unaltered group (Total data points)		
KRAS	5 (4.8%)	217 (20.9%)	-2.13	Mutual exclusivity
STK11	3 (2.9%)	108 (10.4%)	-1.86	Mutual exclusivity
PTEN	5 (4.8%)	62 (5.9%)	-0.33	Mutual exclusivity
PIK3CA	12 (11.4%)	82 (7.9%)	0.53	Co-occurrence
KEAP1	7 (6.7%)	155 (14.9%)	-1.16	Mutual exclusivity
CUL3	5 (4.8%)	34 (3.3%)	0.54	Co-occurrence
EGFR	2 (1.9%)	115 (11.1%)	-2.54	Mutual exclusivity

Table 4.2.4: Co-occurrence of mutations in NFE2L2 with mutations in other genes that are associated with the NRF2-signalling pathway

(A) Eight NFE2L2 mutant lung cancer cell lines were filtered for the presence of mutations in KRAS, STK11, PTEN, PIK3CA, KEAP1, CUL3 and EGFR. Mutations with a positive log ratio are designated as having a tendency of co-occurrence with NFE2L2 mutations, whereas those with a negative log ratio are designated as having a tendency of mutual exclusivity with mutations in NFE2L2. (B) 1144 pan-lung tumour samples were filtered for mutations in NFE2L2, yielding 150 NFE2L2 tumours then further sub-filtered for the presence of mutations in KRAS, STK11, PTEN, PIK3CA, KEAP1, CUL3 and EGFR. Log ratios and tendency of mutations to occur with NFE2L2 mutations were assigned as indicated above.

4.2.5 Mutations in *CUL3* co-occur alongside mutations in *PTEN, PIK3CA* and *NFE2L2* in both lung cancer cell lines and tumours

The co-occurrence of somatic mutations in *CUL3* with other mutations that could influence NRF2 degradation was analysed in both the lung cell line and tumour datasets. First the lung cancer cell lines and tumours were filtered for the presence of mutations in *CUL3*, which identified eight *CUL3* mutant cell lines and 51 *CUL3* mutant tumours. *CUL3* mutant cell lines and tumour samples were then filtered for the presence of additional mutations in *KRAS, STK11, PTEN, PIK3CA, KEAP1, NFE2L2* and *EGFR*.

In the lung cancer cell lines, *CUL3* mutations appear to co-exist with mutations in *KRAS, PTEN, PIK3CA, KEAP1, NFE2L2* and *EGFR*, and are

mutually exclusive with mutations in *STK11* (Table 4.2.5A). These findings either suggest that *CUL3* mutations occur alongside many other mutations that will influence the degradation of NRF2, so *CUL3* mutations themselves may not play a huge role in halting NRF2 degradation, or that the number of *CUL3* mutants analysed is not sufficient to produce robust data. In lung tumours, mutations in *CUL3* co-occur with mutations in *PTEN, PIK3CA*, and *NFE2L2*, and are mutually exclusive with mutations in *KRAS, STK11, KEAP1* and *EGFR*, which is in agreement with previously published data indicating the *CUL3* and *PTEN* mutations complement each other (Dorr et al., 2015). The tumour dataset has a considerably larger number of *CUL3* mutant samples, indicating that these data may be more reliable. Interestingly, *CUL3* mutations occur alongside *PTEN* mutations in both the lung cell line and tumour datasets, whereas *KEAP1* and *NFE2L2* mutations have been shown to be mutually exclusive to mutations in *PTEN* in every analysis.

A

***CUL3* mutant Cell lines**				
Gene	Frequency of alteration		Log ratio	Direction/Tendency
	In altered group (Total *CUL3* mutants)	In unaltered group (Total cell lines)		
KRAS	1 (12.5%)	7 (2.9%)	2.09	Co-occurrence
STK11	0 (0%)	8 (3.4%)	-	Mutual exclusivity
PTEN	1 (12.5%)	7 (2.9%)	2.09	Co-occurrence
PIK3CA	3 (37.5%)	5 (2.1%)	4.17	Co-occurrence
KEAP1	1 (12.5%)	7 (2.9%)	2.09	Co-occurrence
NFE2L2	1 (12.5%)	7 (2.9%)	2.09	Co-occurrence
EGFR	1 (12.5%)	7 (2.9%)	2.09	Co-occurrence

B

CUL3 mutant Tumour samples				
Gene	Frequency of alteration		Log ratio	Direction/ Tendency
	In altered group (Total *CUL3* mutants)	In unaltered group (Total data points)		
KRAS	4 (7.8%)	218 (19.9%)	-1.35	Mutual exclusivity
STK11	3 (5.9%)	108 (9.9%)	-0.75	Mutual exclusivity
PTEN	4 (7.8%)	63 (5.8%)	0.44	Co-occurrence
PIK3CA	6 (11.8%)	88 (8.1%)	0.55	Co-occurrence
KEAP1	6 (11.8%)	156 (14.3%)	-0.28	Mutual exclusivity
NFE2L2	5 (9.8%)	79 (7.2%)	0.44	Co-occurrence
EGFR	3 (5.9%)	114 (10.4%)	-0.83	Mutual exclusivity

Table 4.2.5: Co-occurrence of mutations in CUL3 with mutations in other genes that are associated with the NRF2-signalling pathway

(A) Eight CUL3 mutant lung cancer cell lines were further sub-filtered for the presence of mutations in KRAS, STK11, PTEN, PIK3CA, KEAP1, NFE2L2 and EGFR. Mutations with a positive log ratio are designated as having a tendency to co-occur with CUL3 mutations, whereas those with a negative log ratio are designated as having a tendency of mutual exclusivity with mutations in CUL3. (B) 1144 pan-lung tumour samples were filtered for the expression of CUL3 mutation, yielding 51 CUL3 mutant tumours. Then further sub-filtered for the presence of mutations in KRAS, STK11, PTEN, PIK3CA, KEAP1, NFE2L2 and EGFR. Log ratios and tendency of mutations to occur with CUL3 mutations were assigned as indicated above.

4.2.6 *KEAP1* mutations co-occur with *MET* and *MYC* mutations in both lung cancer cell lines and tumours

Several driver mutations, such as mutations in *EGFR, KRAS, PTEN* and *MET* have been identified in lung cancer (Luo and Lam, 2013; Zhu et al., 2017). Since somatic mutations in *KEAP1* tend to co-occur with oncogenic driver mutations in *KRAS* and *STK11*, we wondered if *KEAP1* mutations co-exist with mutations in other driver genes. We therefore filtered both the lung cancer cell line dataset and the tumour dataset for samples h8arbouring mutations in *KEAP1*, and then screened this *KEAP1* mutant sub-population for the presence of additional mutations in the lung cancer driver genes: *BRAF, CDKN2A, ERBB2,*

HRAS, MET, MYC, NOTCH and *TP53.* This revealed that in both the lung cancer cell lines and the lung tumours, *KEAP1* mutations co-exist with mutations in *HRAS, MET,* and *MYC* (Table 4.2.6). Interestingly, in the cell line data set, *KEAP1* mutations co-exist with mutations in *TP53*, a commonly mutated gene in NSCLC. By contrast, in the tumour data set, *KEAP1* mutations appear to be mutually exclusive with mutations in *TP53.* One potential reason for this discrepancy between the cell line and tumour data sets is that in the process of cancer cell immortalization they often undergo a procedure that alters the status of *TP53* (Metz et al., 1995). Also *TP53* mutation status in human cancer cell lines is quite controversial and often incorrect (Leroy et al., 2014).

A

***KEAP1* mutant cell lines**				
Gene	Frequency of alteration		Log ratio	Direction/Tendency
	In altered group (Total *KEAP1* mutants)	In unaltered group (Total cell lines)		
BRAF	2 (7.1%)	26 (10.9%)	-0.62	Mutual exclusivity
CDKN2A	13 (46.4%)	15 (6.3%)	2.88	Co-occurrence
ERBB2	4 (14.3%)	24 (10.0%)	0.52	Co-occurrence
HRAS	1 (3.6%)	27 (11.3%)	-1.65	Mutual exclusivity
MET	3 (10.7%)	25 (10.5%)	0.03	Co-occurrence
MYC	8 (28.6%)	20 (8.4%)	1.77	Co-occurrence
NOTCH	4 (14.3%)	24 (10.0%)	0.52	Co-occurrence
TP53	27 (96.4%)	1 (0.4%)	7.9	Co-occurrence

B

***KEAP1* mutant tumour Samples**				
Gene	Frequency of alteration		Log ratio	Direction/ Tendency
	In altered group(Total *KEAP1* mutants)	In unaltered group (Total cell lines)		
BRAF	8 (4.7%)	62 (6.4%)	-0.44	Mutual exclusivity
CDKN2A	12 (7.1%)	89 (9.1%)	-0.37	Mutual exclusivity
ERBB2	2 (1.2%)	24 (2.5%)	-1.07	Mutual exclusivity
HRAS	2 (1.2%)	9 (0.9%)	0.35	Co-occurrence
MET	6 (3.5%)	28 (2.9%)	0.30	Co-occurrence
MYC	2 (1.2%)	6 (0.6%)	0.93	Co-occurrence
NOTCH	7 (4.1%)	64 (6.6%)	-0.67	Mutual exclusivity
TP53	103 (60.6%)	672 (68.9%)	-0.19	Mutual exclusivity

Table 4.2.6: Co-occurrence of mutations in KEAP1 with mutations in other genes that are commonly associated with lung cancer.

(A) 28 KEAP1 mutant lung cancer cell lines were filtered for the presence of mutations in BRAF, CDKN2A, ERBB2, HRAS, MET, MYC, NOTCH and TP53. Mutations with a negative log ratio are designated as having a tendency of mutual exclusivity with mutations in KEAP1, whereas those with a positive log ratio are designated as having a tendency of co-occurrence with KEAP1 mutations. (B)1144 pan-lung cancer tumours were filtered for the presence of mutations in KEAP1, yielding 170 KEAP1 mutant tumours. Then further sub-filtered for the presence of mutations in: BRAF, CDKN2A, ERBB2, HRAS, MET, MYC, NOTCH and TP53. Log ratios and tendency of mutations to occur with KEAP1 mutations were assigned as described above.

4.2.7 Mutations in *NFE2L2* co-occur with mutations *in MET, MYC, NOTCH* and *TP53*

In view of the fact that somatic mutations in *NFE2L2* do not occur with mutations in the oncogenic driver gene *KRAS* (Table 4.2.4), we examined whether *NFE2L2* mutations might co-occur with other common oncogenic mutations. We therefore first filtered both the lung cancer cell line dataset and the tumour dataset for the presence of a *NFE2L2* mutation, and then screened the two *NFE2L2* mutant sub-populations for additional mutations in *BRAF, CDKN2A, ERBB2, HRAS, MET, MYC, NOTCH* and *TP53.* Surprisingly, *NFE2L2* mutations were found to co-occur with mutations in *TP53* in both the cancer cell line and tumour datasets (Table 4.2.7). We surmise that because mutations resulting in NRF2 upregulation seem to be involved in cancer progression rather than cancer initiation, mutations in *NFE2L2 and KEAP1* will co-exist with mutations in oncogenic driver genes. Thus collectively, our analyses indicate that *KEAP1* mutations co-exist with mutations in *KRAS*, whereas *NFE2L2* mutations co-exist with mutations in *TP53.*

A

NFE2L2 mutant cell lines				
Gene	Frequency of alteration		Log ratio	Direction/Tendency
	In altered group (Total *NFE2L2* mutants)	In unaltered group (Total cell lines)		
BRAF	0 (0%)	8 (3.3%)	-	Mutual exclusivity
CDKN2A	2 (25%)	6 (2.5%)	3.32	Co-occurrence
ERBB2	0 (0%)	8 (3.3%)	-	Mutual exclusivity
HRAS	0 (0%)	8 (3.3%)	-	Mutual exclusivity
MET	1 (12.5%)	7 (2.9%)	2.11	Co-occurrence
MYC	1 (12.5%)	7 (2.9%)	2.11	Co-occurrence
NOTCH	1 (12.5%)	7 (2.9%)	2.11	Co-occurrence
TP53	8 (100%)	0 (0%)	-	Co-occurrence

B

NFE2L2 mutant tumour samples				
Gene	Frequency of alteration		Log ratio	Direction/ Tendency
	In altered group (Total *NFE2L2* mutants	In unaltered group (Total data points)		
BRAF	5 (4.8%)	65 (6.3%)	-0.39	Mutual exclusivity
CDKN2A	20 (19.1%)	81 (7.8%)	1.29	Co-occurrence
ERBB2	1 (0.9%)	25 (2.4%)	-1.34	Mutual exclusivity
HRAS	1 (0.9%)	10 (0.9%)	-0.02	Mutual exclusivity
MET	5 (4.8%)	29 (2.8%)	0.77	Co-occurrence
MYC	0 (0.00%)	8 (0.8%)	<-10	Mutual exclusivity
NOTCH	8 (7.6%)	63 (6.1%)	0.33	Co-occurrence
TP53	88 (83.8%)	687 (66.1%)	0.34	Co-occurrence

Table 4.2.7: Co-occurrence of mutations in NFE2L2 with mutations in other genes that are commonly associated with lung cancer.

(A) Eight NFE2L2 mutant lung cancer cell lines were sub-filtered for the presence of mutations in BRAF, CDKN2A, ERBB2, HRAS, MET, MYC, NOTCH and TP53. Mutations with a negative log ratio are designated as having a tendency of mutual exclusivity with mutations in NFE2L2, whereas those with a positive log ratio are designated as having a tendency of co-occurrence with NFE2L2 mutations. (B)1144 pan-lung cancer tumours were filtered for the presence of mutations in NFE2L2, yielding 150 NFE2L2 mutant tumours. Which were then further sub-filtered for the presence of mutations in BRAF, CDKN2A, ERBB2, HRAS, MET, MYC, NOTCH and TP53. Log ratios and tendency of mutations to occur with NFE2L2 mutations were assigned as indicated above.

4.2.8 *CUL3* mutations co-occur with mutations in *MYC, NOTCH and TP53,* in both lung cancer cell lines and tumours

Since somatic mutations in *CUL3* show mutual exclusivity with *KRAS* mutations, we wondered whether they co-exist with mutations in *TP53*, as is the case for somatic mutations in *NFE2L2*, or with other oncogenic driver genes. Both the cancer cell line dataset and the tumour dataset were filtered for the presence of mutations in *CUL3* and this *CUL3* mutant sub-population was then screened for additional mutations in: *BRAF, CDKN2A, ERBB2, HRAS, MET, MYC, NOTCH* and *TP53.* Interestingly, like *NFE2L2* mutants, the cancer cell lines and the tumour samples with mutant *CUL3* both co-existed with mutations in *TP53* (Table 4.2.8). This is possibly expected because our previous analysis revealed that mutations in *NFE2L2* and *CUL3* can co-occur.

A

***CUL3* mutant cell lines**				
Gene	Frequency of alteration		Log ratio	Direction/Tendency
	In altered group (Total *CUL3* mutants)	In unaltered group (Total cell lines)		
BRAF	0 (0%)	8 (3.4%)	-	Mutual exclusivity
CDKN2A	1 (12.5%)	7 (2.9%)	2.09	Co-occurrence
ERBB2	1 (12.5%)	7 (2.9%)	2.09	Co-occurrence
HRAS	0 (0%)	8 (3.4%)	-	Mutual exclusivity
MET	3 (37.5%)	5 (2.1%)	4.16	Co-occurrence
MYC	4 (50%)	4 (1.7%)	4.90	Co-occurrence
NOTCH	1 (12.5%)	7 (2.9%)	2.09	Co-occurrence
TP53	6 (75%)	2 (0.8%)	6.48	Co-occurrence

B

CUL3 mutant tumour Samples				
Gene	Frequency of alteration		Log ratio	Direction/ Tendency
	In altered group (Total *CUL3* mutants)	In unaltered group (Total cell lines)		
BRAF	5 (9.8%)	65 (5.9%)	0.72	Co-occurrence
CDKN2A	3 (5.9%)	98 (8.9%)	-0.61	Mutual exclusivity
ERBB2	1 (1.9%)	25 (2.3%)	-0.22	Mutual exclusivity
HRAS	1 (1.9%)	10 (0.9%)	1.10	Co-occurrence
MET	1 (1.9%)	33 (3.02%)	-0.62	Mutual exclusivity
MYC	1 (1.9%)	7 (0.6%)	1.61	Co-occurrence
NOTCH	4 (7.8%)	67 (6.1%)	0.36	Co-occurrence
TP53	42 (82.4%)	733 (67.1%)	0.30	Co-occurrence

Table 4.2.8: Co-occurrence of mutations in CUL3 with mutations in other genes that are commonly associated with lung cancer

(A) Eight CUL3 mutant lung cancer cell lines were sub-filtered for the presence of mutations in BRAF, CDKN2A, ERBB2, HRAS, MET, MYC, NOTCH and TP53. Mutations with a negative log ratio are designated as having a tendency of mutual exclusivity with mutations in CUL3, whereas those with a positive log ratio are designated as having a tendency of co-occurrence with CUL3 mutations. (B) 1144 pan-lung cancer tumours 51 were filtered for the presence of mutations in CUL3, then further sub-filtered for the presence of mutations in BRAF, CDKN2A, ERBB2, HRAS, MET, MYC, NOTCH and TP53. Log ratios and tendency of mutations to occur with CUL3 mutations were assigned as described above.

4.2.9 Most *KEAP1*-mutant cancer cell lines harbour homozygous missense mutations that translate into single amino acid changes distributed across the protein

To gain greater understanding of the impact that somatic mutations in *KEAP1* have in terms of impairing NRF2-degradation we analysed the zygosity and the location of the mutations in the lung cancer cell lines. This revealed that amongst the 28 *KEAP1*-mutant cell lines in the cancer cell dataset (see, Table 4.2.1), 89% of them processes homozygous mutations. The majority of the mutations identified were single missense mutations, which often lead to the introduction of a Cys amino acid in the protein (Table 4.2.9). Consistent with previously published data, we also saw that none of the key redox sensitive Cys residues were mutated (Hayes and McMahon, 2009). Also, the majority of *KEAP1*

mutations were found in ADC cell lines, which supports data generated by the TCGA consortium showing that *KEAP1* mutations are more prevalent in ADC than SQCC (The Cancer Genome Atlas Research, 2014).

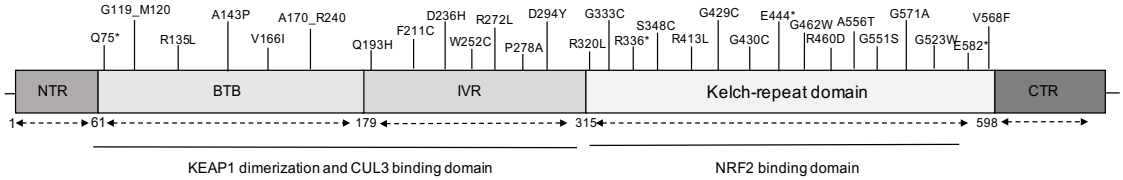
In an attempt to understand the significance of the somatic mutations in *KEAP1*, we mapped them onto a schematic of the protein. Like others (Hast et al., 2014), we found that mutations in *KEAP1* translate into changes throughout the entire protein, showing no clear focal point (Figure 4.2.1A). This mutational pattern is not only seen in lung cancer, *KEAP1* is frequently mutated in head and neck cancer also showing no clear focal point of mutation (The Cancer Genome Atlas et al., 2015). Quantification of the localization of the mutations suggests that a disproportionate number of mutations occur in regions of *KEAP1* encoding the Kelch-repeat domain (Figure 4.2.1B). This is to be expected as the Kelch-repeat domain is the region of KEAP1 that binds NRF2, and mutations in this site could potentially lead to increased NRF2 activation through inhibition of binding and subsequent degradation of NRF2. Furthermore, mutations in both the IVR and BTB domains of KEAP1 were also identified, which will potentially effect KEAP1 dimerization and CUL3 binding.

Cell line	Disease subtype	Location of mutation	Heterozygous / Homozygous	Variant classification
A549	ADC	p.G333C	Homozygous	Missense mutation
BEN	Lung cancer	p.A556T	Homozygous	Missense mutation
HCC44	ADC	p.F211C	Homozygous	Missense mutation
NCIH1355	ADC	p.Q75*	Homozygous	Nonsense mutation
NCIH1417	Small cell carcinoma	p.P278A	Heterozygous	Missense mutation
NCIH1435	ADC	p.R413L	Homozygous	Missense mutation
NCIH1573	ADC	p.A143P	Homozygous	Missense mutation
NCIH1623	ADC	p.R320L	Homozygous	Missense mutation
NCIH1688	Small cell carcinoma	p.G511S	Homozygous	Missense mutation
NCIH1755	ADC	p.E582*	Homozygous	Nonsense mutation
NCIH1792	ADC	p.G462W	Homozygous	Missense mutation
NCIH1915	Large cell carcinoma	p.R135L	Homozygous	Missense mutation
NCIH1944	ADC	p.W252C	Homozygous	Missense mutation
NCIH2023	ADC	p.V568F	Homozygous	Missense mutation
NCIH2110	NSCLC	p.G429C	Homozygous	Missense mutation
NCIH2122	ADC	p.A170_R204 deletion	Heterozygous	Deletion in frame
NCIH2170	SQCC	p.R336*	Homozygous	Nonsense mutation
NCIH2172	NSCLC	p.G430C	Homozygous	Missense mutation
NCIH23	ADC	p.Q193H	Homozygous	Missense mutation
NCIH322M	ADC	p.R460D	Homozygous	Missense mutation
NCIH460	Large cell carcinoma	p.D236H	Homozygous	Missense mutation
NCIH647	SQCC	p.G523W	Homozygous	Missense mutation
NCIH661	Large cell carcinoma	p.V168I	Homozygous	Missense mutation
NCIH740	Small cell carcinoma	p.S348C	Homozygous	Missense mutation
NCIH838	ADC	p.E444*	Homozygous	Nonsense mutation
RERFLCMS	ADC	p.G119_M120	Heterozygous	Deletion in frame
SHP77	Small cell carcinoma	p.D294Y	Homozygous	Missense mutation
UMC11	Lung carcinoid-endocrine tumour	p.G571A	Homozygous	Missense mutation

Table 4.2.9: Most KEAP1 mutant cell lines harbor homozygous missense mutations and are of the adenocarcinoma histological subtype

Table showing the name of the 28 cell lines harboring mutations in KEAP1 and the disease subtype, location that the mutation occurs in the protein, whether it is heterozygous/homozygous and the variant classification. An asterisk (*) symbol indicates the insertion of a stop codon. The p. symbol indicates a molecular change in the primary structure of the protein.

A



B

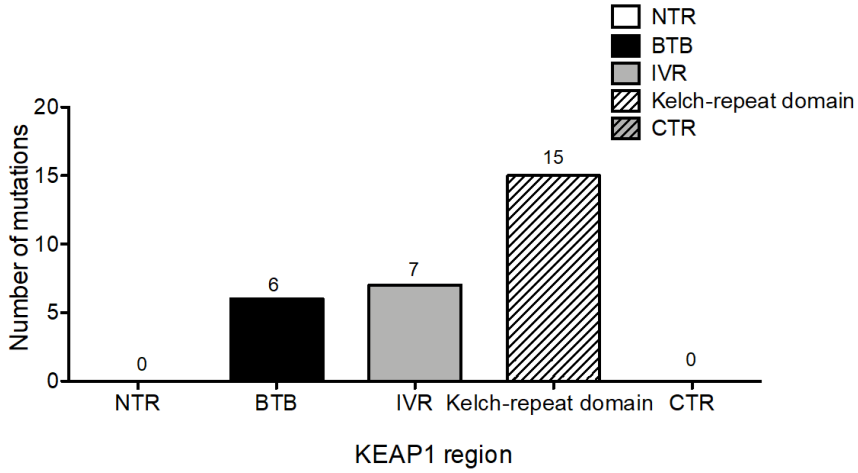


Figure 4.2.1: Most KEAP1 mutant cell lines harbor mutations that are distributed throughout the protein structure

(A) Mapping of the KEAP1 mutations found in the 28 cell lines onto a schematic of the KEAP1 protein structure. The locations of each mutations are indicated above the schematic. The amino acid numbering conferring to each specific functional domain of the protein is given below separated by dotted arrows. (B) Graphical quantification of (A).

4.2.10 Most *NFE2L2* mutant cell lines harbour heterozygous missense mutations, located in the Neh2 domain

To gain an understanding of the impact that somatic mutations in *NFE2L2* have in terms of NRF2 function we analysed the zygosity and location of the mutations in the lung cancer cell lines. Out of the eight *NFE2L2* mutant cell lines (from Table 4.2.1), all but one cell line, LK-2, possess heterozygous mutations (Table 4.2.10). The majority of these cell lines have missense mutations in *NFE2L2* and were from a range of diverse histological backgrounds. When we mapped these mutations onto a schematic of the NRF2 protein, the majority (75%) localized to the Neh2 domain (Figure 4.2.2). However, we did identify one cell line that harboured a mutation in the Neh5 domain and one in the Neh1 domain. Mutations in the Neh2 domain have been documented to lead to increased NRF2 activation through halting KEAP1-dependent degradation of the transcription factor. In the case of the Neh5 and Neh1 domains, they are involved in transactivation and DNA-binding, respectively, and could therefore also influence NRF2 activity in tumours. Interestingly, MED16 has been shown to bind directly to the Neh5 and Neh1 domains in NRF2 and positions the transcription factor close to the transcriptional machinery (Sekine et al., 2015). MED16 interaction the NRF2 has been demonstrated to be essential for the inducible transcription of several NRF2-target genes. Thus, mutations in both the Neh1 and Neh5 could potentially alter NRF2 activity, possibly via MED16, but it is difficult to determine what effect a mutation in the Neh5 or Neh1 domain would have on NRF2 degradation.

NFE2L2 is often also mutated in head and neck cancer with a tendency for mutations to occur in the Neh2 domain but also a few mutations in Neh5, Neh7 and Neh1 domains (The Cancer Genome Atlas et al., 2015). Interestingly *NFE2L2* mutations in both lung cancer and head and neck cancer do not occur in the Neh6 domain, which is the site of GSK-3 phosphorylation and β-TrCP binding.

Cell line	Disease subtype	Location of mutation	Heterozygous / Homozygous mutation	Variant classification
CORL88	Small cell carcinoma	p.Q26P	Heterozygous	Missense mutation
EBC1	SQCC	p.D77V	Heterozygous	Missense mutation
LK2	NSCLC	p.E79K	Homozygous	Missense mutation
NCIH1436	Small cell carcinoma	p.A496V	Heterozygous	Missense mutation
NCIH1568	ADC	p.D77_E79 deletion	Heterozygous	Deletion in frame
NCIH2228	NSCLC	p.G31A	Heterozygous	Missense mutation
NCIH748	Small cell carcinoma	p.V175F	Heterozygous	Missense mutation
RERFLCSQ1	SQCC	p.D29H	Heterozygous	Missense mutation

Table 4.2.10: Most NFE2L2 mutant cell lines harbor heterozygous missense mutations and are from various histological backgrounds

Table showing the name of the eight cell lines harboring mutations in NFE2L2 and the disease subtype, location of the mutation, whether it is heterozygous / homozygous and the variant classification.

A

B

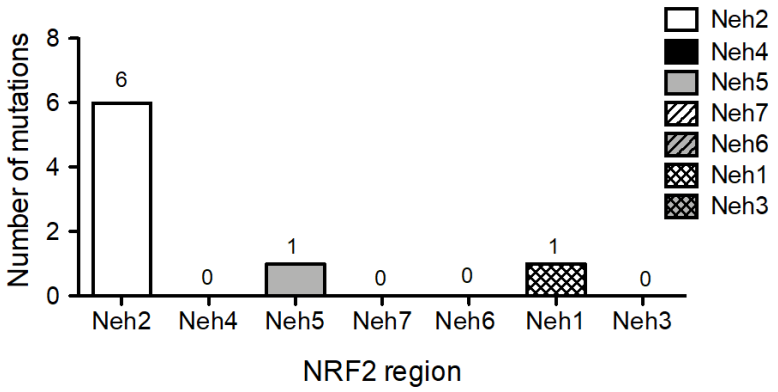


Figure 4.2.2: Most NFE2L2 mutant cell lines harbor mutations in the gene that encode the Neh2 domain of NRF2

(A) Mapping of the NFE2L2 mutations found in the eight cell lines onto a schematic of NRF2 protein structure. The location of each mutation are indicated above the schematic. The amino acid numbering conferring to each specific functional domain of the protein is given below separated by dotted arrows. (B) Graphical quantification of (A).

4.2.11 Most *CUL3* mutant cell lines harbour heterozygous missense mutations, located in the first half of the *CUL3* protein

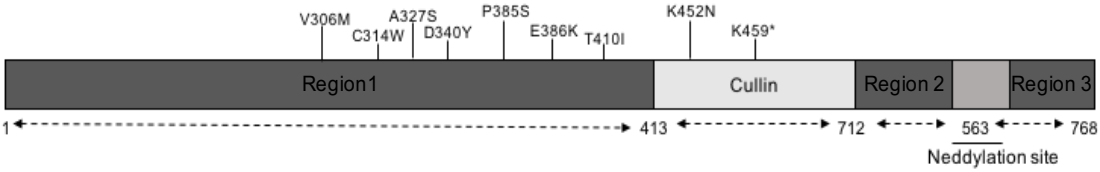
To gain greater understanding of the impact that somatic mutations in *CUL3* have in terms of impairing NRF2-degradation we analysed the zygosity and location of the mutations in the lung cancer cell lines. Table 4.2.11 shows that the majority, (88.9%), of *CUL3* mutant cell lines possess heterozygous missense mutations and were from a range of diverse histological backgrounds. In an attempt to ascribe a function to the *CUL3* mutations, we mapped them onto a schematic of the CUL3 protein. This revealed that 77.8% of mutations were located between amino acids 1- 413, a region that we have referred to as region 1 (Figure 4.2.3); the N-terminal region 1 is thought to be the site that binds the BTB subunit of KEAP1 (Canning et al., 2015). We also identified two mutations in the cullin site. It is difficult determine what effect theses mutations will have on NRF2 degradation.

Cell line	Disease subtype	Location of mutation	Heterozygous / Homozygous mutation	Variant classification
DMS53	Small cell carcinoma	p.A327S	Homozygous	Missense mutation
RERFLCSQ1	SQCC	p.C314W	Heterozygous	Missense mutation
NCIH2342	NSCLC	p.D340Y	Heterozygous	Missense mutation
LC2AD	ADC	p.E386K	Heterozygous	Missense mutation
NCIH2081	Small cell carcinoma	p.K452N	Heterozygous	Missense mutation
NCIH1793	ADC	p.V306M/p.K459*	Heterozygous /Heterozygous	Missense/ nonsense mutation
NCIH1105	Small cell carincoma	p.P385S	Heterozygous	Missense mutation
NCIH460	Large cell carcinoma	p.T410I	Heterozygous	Missense mutation

Table 4.2.11*: *Most CUL3 mutant cell lines harbor heterozygous missense mutations

Table showing the name of the eight cell lines harboring mutations in CUL3 and the disease subtype, location of the mutation, whether it is heterozygous / homozygous and the variant classification

A



B

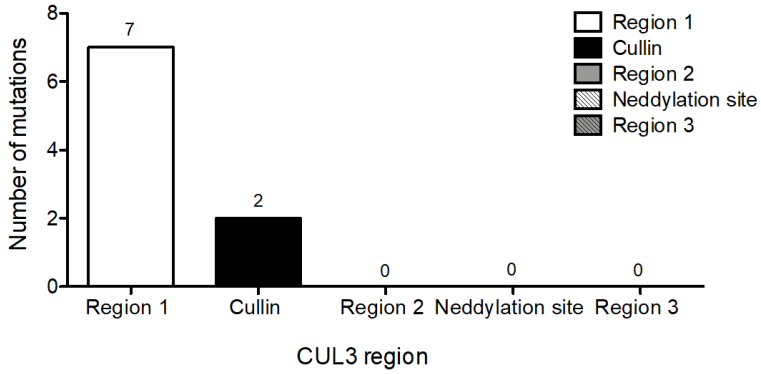


Figure 4.2.3: Most CUL3 mutant cell lines harbor mutations that map to region 1 of the protein

(A) Mapping of the CUL3 mutations found in the eight cell lines onto the surface of a schematic of the CUL3 protein. The location of each mutation is indicated above the schematic. The amino acid numbering conferring to each specific functional domain of the protein is given below separated by dotted arrows. (B) Graphical quantification of (A).

4.2.12 *KEAP1* mutant cell lines with the highest mRNA expression of *NQO1* tend to have mutations in the Kelch-repeat domain

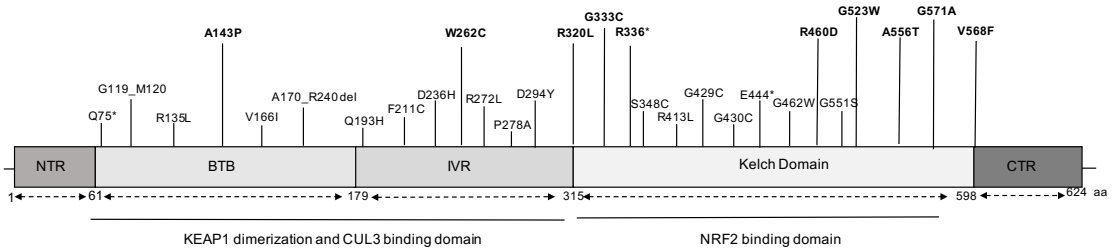
To further understand what effect *KEAP1* mutations have on NRF2 degradation, we analysed expression of the NRF2-target gene *NQO1* in the cell lines harbouring *KEAP1* mutations. Table 4.2.12 shows that out of the 28 *KEAP1*-mut cell lines, those with the highest *NQO1* mRNA levels all harbour homozygous missense mutations. Suggesting that a homozygous mutation in *KEAP1* leads to a greater increase in NRF2 activity than a heterozygous mutation. Figure 4.2.4 shows that when the locations of the *KEAP1* mutations in these cell lines were mapped onto the structure of KEAP1 protein, the majority of mutations (80%) are located in regions of the gene that encode the Kelch-repeat domain. The same analysis was carried out looking at NRF2 mRNA levels (data not shown). This in combination with the data shown in section 4.2.9, indicates that the majority of *KEAP1* mutations found in lung cancer cell lines occur in the Kelch-repeat domain and mutations in this region will inhibit NRF2 degradation and subsequently enhance NRF2-target gene expression. Collectively, these results suggest that the presence of a *KEAP1* mutation in lung cancer is an indicator of enhanced NRF2 activity.

Top ten *KEAP1* mutant cell lines with the highest expression of NQO1 mRNA				
Cell line	**Disease subtype**	**Location of mutation**	**Heterozygous / Homozygous mutation**	**Variant classification**
NCIH1573	ADC	p.A143P	Homozygous	Missense mutation
NCIH1944	ADC	p.W252C	Homozygous	Missense mutation
NCIH1623	ADC	p.R320L	Homozygous	Missense mutation
A549	ADC	p.G333C	Homozygous	Missense mutation
NCIH2170	SQCC	p.R336*	Homozygous	Nonsense mutation
NCIH322M	ADC	p.R460D	Homozygous	Missense mutation
NCIH647	AQCC	p.G523W	Homozygous	Missense mutation
BEN	Lung cancer	p.A556T	Homozygous	Missense mutation
UMC11	Lung carcinoid-endocrine tumour	p.G571A	Homozygous	Missense mutation
NCIH2023	ADC	p.V568F	Homozygous	Missense mutation

Table 4.2.12: Characteristics of the KEAP1 mutant cell lines with the highest expression of NQO1 mRNA

Table highlighting the top ten KEAP1 mutant cell lines with the highest expression of NQO1 mRNA, in no particular order. The name of the cell lines are given in the far left-hand side column, then the disease subtype, location of the mutation, the zygosity of the mutation and the variant classification; as indicated in the second top line of the table.

A



B

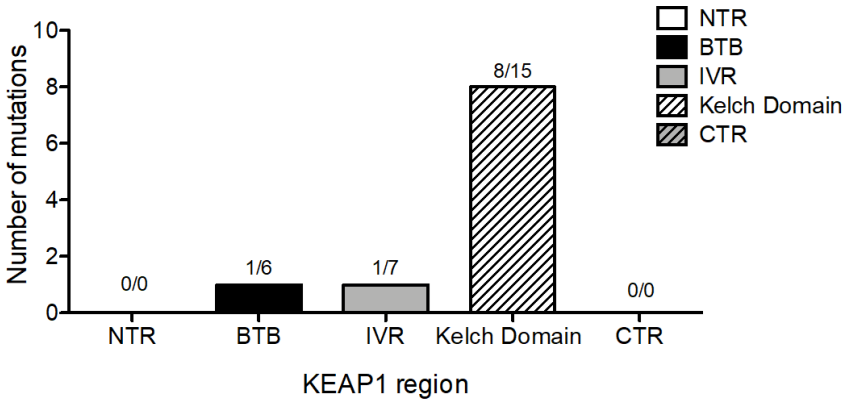


Figure 4.2.4: The KEAP1 mutant cell lines with the highest expression of NQO1 predominantly possess mutations in the Kelch-repeat domain

(A) The locations of the mutations associated with high levels of NQO1 are highlighted in bold font and are elevated above the schematic representing the KEAP1 protein. (B) Graphical quantification of (A) showing that the top ten NQO1 expressing KEAP1 mutant cell lines. The first number displayed above the bars highlights the number of mutations found in that specific KEAP1 region that are found in the top ten NQO1 expressing cell lines. The number after represents the total number of mutations found in that region of KEAP1 out of the 28 cell lines analysed.

4.2.13 The *NFE2L2* mutant cell lines that express highest levels of *NQO1* only have mutations in the Neh2 domain of NRF2

To further understand what effect *NFE2L2* mutations have on NRF2 activity, we analysed the expression of the NRF2-target gene *NQO1* in cell lines harbouring *NFE2L2* mutations. Table 4.2.13 and Figure 4.2.5 show that out of the eight *NFE2L2*-mut cell lines, those with the highest *NQO1* expression predominantly harbour heterozygous missense mutations in the Neh2 domain of NRF2, more specifically localized to the KEAP1 binding regions, DLG and ETGE. MED16 interaction with NRF2 has been shown to be critical for the transcription of the NRF2-target gene, *NQO1* (Sekine et al., 2015). This could be impaired in NCIH748 and NCIH1436 cells which harbour mutations in the Neh5 and Neh1 domains respectively, accounting for their low NQO1 expression. Taken together with the data in section 4.2.10, these findings indicate that *NFE2L2* mutations predominantly occur in the Neh2 domain of the protein and will lead to enhanced NRF2-target gene expression through inhibiting the binding of KEAP1 to NRF2 and subsequent proteasomal degradation of the transcription factor.

Top five *NFE2L2* mutant cell lines with the highest expression of NQO1 mRNA				
Cell line	**Disease subtype**	**Location of mutation**	**Heterozygous / Homozygous mutation**	**Variant classification**
EBC1	SQCC	p.D77V	Heterozygous	Missense mutation
LK2	NSCLC	p.E79K	Homozygous	Missense mutation
NCIH1436	Small cell carcinoma	p.A496V	Heterozygous	Missense mutation
NCIH2228	NSCLC	p.G31A	Heterozygous	Missense mutation
RERFLCSQ1	SQCC	p.D29H	Heterozygous	Missense mutation

Table 4.2.13: The top five NFE2L2 mutant cell lines with the highest expression of NQO1 mRNA

Table highlighting the top five NFE2L2 mutant cell lines with the highest expression of NQO1 mRNA, in no particular order. The name of the cell lines are given in the far left-hand side column, then the disease subtype, location of the mutation, the zygosity of the mutation and the variant classification, as indicated in the second top line of the table.

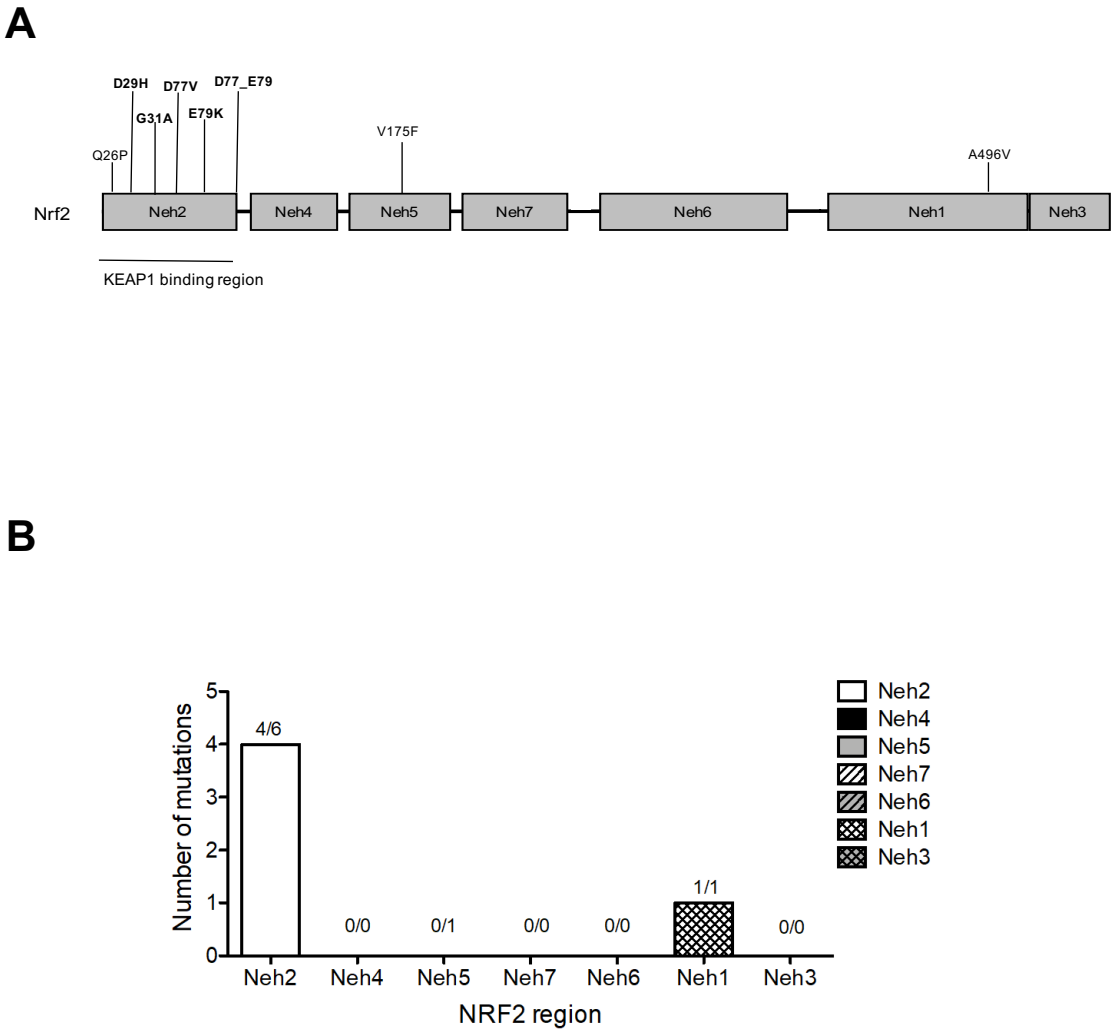


Figure 4.2.5: The NFE2L2 mutant cell lines with highest expression of NQO1 mRNA possess mutations in the Neh2 region of NRF2

(A) The locations of the mutations associated with high levels of NQO1 are highlighted in bold font above the schematic representing the NRF2 protein. (B) Graphical quantification of (A) showing that the top five NQO1 expressing NFE2L2 mutant cell lines. The first number displayed above the bars highlights the number of mutations found in that specific NRF2 region that are found in the top five NQO1 expressing cell lines. The number after represents the total number of mutations found in that region of NFE2L2 out of the eight cell lines analysed.

4.2.14 *CUL3* mutant cell lines with the highest expression of *NQO1* mRNA poses mutations in region 1 of the protein

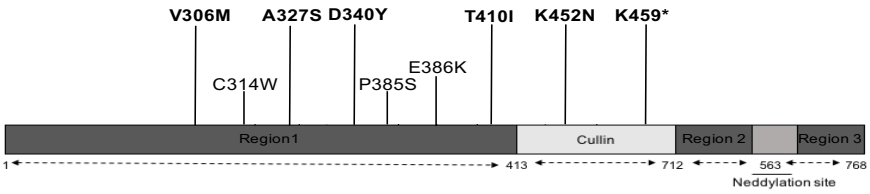
To further understand what effect *CUL3* mutations have on NRF2 degradation, we analysed the expression of the NRF2-target *NQO1* in cell lines harbouring *CUL3* mutations. Table 4.2.14 shows that out of the eight *CUL3*-mut cell lines the top five mutants with the highest expression of *NQO1* process homozygous missense mutations. Figure 4.2.6 shows that the mutations in these *CUL3* mutant cell lines are predominately located in region 1 and few have mutations in the cullin domain. These data, in combination with the data in section 4.2.11, suggest that the majority of mutations in *CUL3* will lead to impaired NRF2 degradation as they occur in region 1 of CUL3 and lead to enhanced *NQO1* expression. This supports the idea that *CUL3* mutations tend to result in a loss of function of the protein (Ooi et al., 2013).

Top five *CUL3* mutant cell lines with the highest expression of NQO1 mRNA				
Cell line	**Disease subtype**	**Location of mutation**	**Heterozygous / Homozygous mutation**	**Variant classification**
DMS53	Small cell carcinoma	p.A327S	Homozygous	Missense mutation
NCIH2342	NSCLC	p.D340Y	Heterozygous	Missense mutation
NCIH2081	Small cell carcinoma	p.K452N	Heterozygous	Missense mutation
NCIH1793	ADC	p.V306M/p.K459*	Heterozygous/Heterozygous	Missense/nonsense mutation
NCIH460	Large cell	p.T410I	Heterozygous	Missense mutation

Table 4.2.14: The top five CUL3 mutant cell lines with the highest expression of NQO1 mRNA

Table highlighting the top five CUL3 mutant cell lines with the highest expression of NQO1 mRNA, in no particular order. The name of the cell lines is given in the far left-hand side column, then the disease subtype, location of the mutation, the zygosity of the mutation and the variant classification, as indicated in the second top line of the table.

A



B

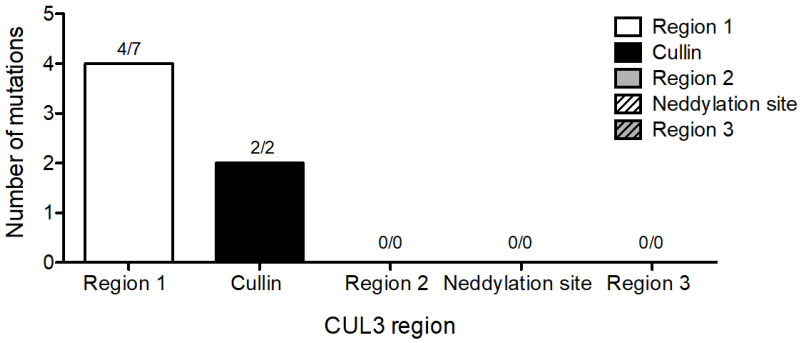


Figure 4.2.6: The top five CUL3 mutant cell lines with the highest expression of NQO1 mRNA tend to have mutations in region 1 of the protein.

(A) The location of the mutations associated with high levels of NQO1 are highlighted in bold font above the schematic representing the CUL3 protein. (B) Graphical quantification of (A) showing that the top five NQO1 expressing CUL3 mutant cell lines. One of the cell lines indicated has a mutation the Cullin region of CUL3 also harbours a mutation in Region 2. The first number displayed above the bars highlights the number of mutations found in that specific CUL3 region that are found in the top five NQO1 expressing cell lines. The number after represents the total number of mutations found in that region of CUL3 out of the eight cell lines analysed.

4.2.15 *KEAP1* mutant cell lines show greater expression of *ABCB6, AKR1B10, AKR1C1, AKR1C2, AKR1C3* and *NQO1*

To determine what effect the presence of a *KEAP1, NFE2L2* or *CUL3* mutation has on the expression of NRF2-target genes associated with drug detoxification, we analysed the expression of *ABCB6, AKR1B10, AKR1C1, AKR1C2, AKR1C3* and *NQO1,* in the presence of each mutation. Using the cell line data set (239 cell lines in total), which has been previously described, we filtered for the presence of either a *KEAP1, NFE2L2* or *CUL3* mutation. We then analysed the expression of six NRF2–target genes associated with drug detoxification in the mutant subgroups for that particular gene against their wildtype counterparts.

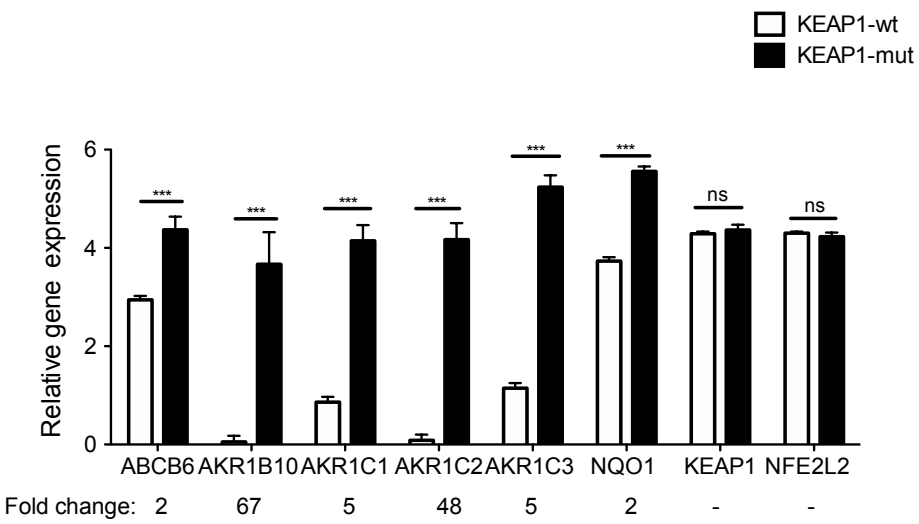
As shown in Figure 4.2.7A, *KEAP1*-mut cell lines (28 cell lines) have significantly higher expression of all of the six target genes analysed in comparison to the *KEAP1*-wt cell lines (211 cell lines). With the most pronounced increase being seen in the expression of *AKR1B10*, which is 67-fold higher in the *KEAP1*-mut cell lines in comparison to the *KEAP1*-wt cell lines.

The expression of the same six target genes was analysed in cell lines harbouring *NFE2L2* mutations and is displayed in Figure 4.2.7B. *NFE2L2*-mut cell lines (8 cell lines) have significantly higher expression of *AKR1C1, AKR1C2, AKR1C3* and *NQO1*, in comparison to *NFE2L2*-wt cell lines (231 cell lines). However, *ABCB6* and *AKR1B10* expression is not significantly altered between the two cell line populations. In comparison to the *KEAP1*-mut cell lines the magnitude of elevation in NRF2-target gene expression in the *NFE2L2*-mut cell lines are considerably less.

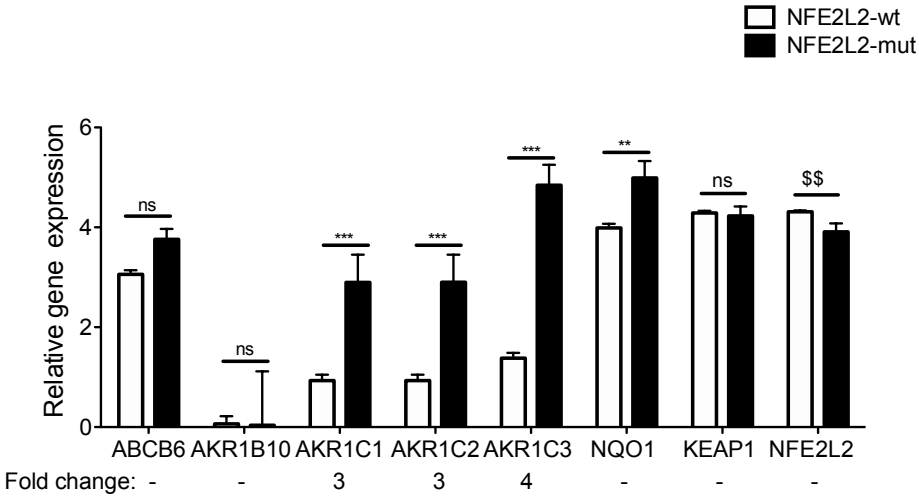
A similar pattern of NRF2-target gene expression as seen in the *NFE2L2*-mut cell lines was seen in the *CUL3*-mut cell lines (8 cell lines) (Figure 4.2.7C), showing significantly increased expression of *AKR1C1, AKR1C2, AKR1C3* and *NQO1* but not significantly altered *ABCB6* or *AKR1B10* expression, in comparison to *CUL3*-wt cell lines (231 cell lines).

These data suggest that the presence of a *KEAP1* mutation has a greater effect on the expression of NRF2-target genes that are involved in drug detoxification, than the presence of a *NFE2L2* or *CUL3* mutation.

A


KEAP1-wt
KEAP1-mut
Relative gene expression
0
2
4
6
***
***
***
***
***
***
ns
ns
ABCB6
AKR1B10
AKR1C1
AKR1C2
AKR1C3
NQO1
KEAP1
NFE2L2
Fold change: 2 67 5 48 5 2 - -

B


NFE2L2-wt
NFE2L2-mut
Relative gene expression
0
2
4
6
ns
ns
***
***
***
**
ns
$$
ABCB6
AKR1B10
AKR1C1
AKR1C2
AKR1C3
NQO1
KEAP1
NFE2L2
Fold change: - - 3 3 4 - - -

C

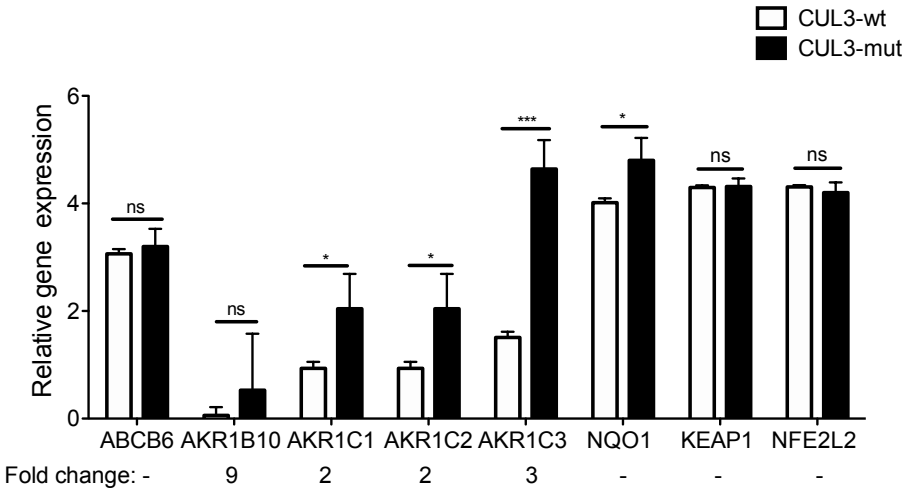

CUL3-wt
CUL3-mut
Relative gene expression
0
2
4
6
ns
ns
*
*
***
*
ns
ns
ABCB6
AKR1B10
AKR1C1
AKR1C2
AKR1C3
NQO1
KEAP1
NFE2L2
Fold change: - 9 2 2 3 - - -

Figure 4.2.7: KEAP1-mut cells show higher expression of all six NRF2-target target genes associated with drug detoxification in comparison to cell lines harboring mutations in NFE2L2 or CUL3

*The white bars on each of the graphs represents the cell line population that is wildtype (wt) for the gene in question and the black bars represents the cell line population that is mutant (mut) for the gene in question. Fold change values for each gene relative to the wt cell line group is displayed under each graph. (A) Analysis of the expression six NRF2-target genes associated with drug detoxification in cell lines that are either KEAP1-mut (28 cell lines) or KEAP1-wt (211 cell lines). The KEAP1-wt cell line population includes cell lines that harbor mutations in NFE2L2 and CUL3. No KEAP1-mut cell line also has a NFE2L2 mutation but one KEAP1-mut cell line has an additional CUL3 mutation. (B) Analysis of the expression of NRF2-target genes associated with drug detoxification in cell lines that are either NFE2L2-mut (8 cell lines) or NFE2L2-wt (231 cell lines). The NFE2L2-wt cell line population includes cell lines that harbor mutations in KEAP1 and CUL3. One NFE2L2-mut cell line also harbors a mutation in CUL3. (C) Analysis of the expression of NRF2-target genes associated with drug detoxification in cell lines that are either CUL3-mut (8 cell lines) or CUL3-wt (231 cell lines). The CUL3-wt cell line population includes cell lines that harbour mutations in KEAP1 and NFE2L2. One CUL3-mut cell line also harbours a mutation in KEAP1 and a separate CUL3 mutant cell line has an additional mutation in NFE2L2. Bars on all three graphs represent the mean value obtained, and the error plotted is the associated SEM. All statistical analysis was carried out using GraphPad Prism 5 software. All significant increases are denoted with the * symbol; * represented a P< 0.05, ** P<0.01 and ** P< 0.001. P values > 0.05 were deemed not significant and are denoted by "ns".*

4.2.16 *KEAP1* mutant cell lines have significantly higher expression of NRF2-target genes associated with glutathione production and utilization, than cell lines harbouring mutations in *NFE2L2* or *CUL3*

To determine if other subgroups of NRF2-target genes show higher expression in *KEAP1* mutant cell lines than *NFE2L2* or *CUL3* mutant cell lines, we analysed the expression of NRF2-target genes associated with glutathione biosynthesis and utilization. As shown in Figure 4.2.8, *KEAP1*-mut cell lines (28 cell lines) show significantly higher expression of *GCLC, GCLM, GPX2* and *SLC7A11* in comparison to *KEAP1*-wt cell lines (211 cell lines), with the highest fold change between *KEAP1*-mut and *KEAP1*-wt cells being in *GPX2* expression, where a 3-fold change was observed. When we analysed the expression of the same four NRF2-target genes in the *NFE2L2*-mut cell lines (8 cell lines) no significant alterations in the expression of *GCLC, GCLM* or *SLC7A11* were

observed, between the *NFE2L2*-mut and *NFE2L2*-wt (230 cell lines) cell line populations. However, a significant increase in the expression of *GPX2* was noted, with a 4-fold change between the mut and wt cell populations. *CUL3*-mut cell lines (8 cell lines) have significantly elevated expression of *GCLM* and *GPX2* in comparison to *CUL3*-wt cell lines (230 cell lines). Also showing the largest fold change in expression in the gene *GPX2*, with a fold change value of 3. The expression of *GCLC* and *SLC7A11* was not significantly altered by the presence of a *CUL3* mutation.

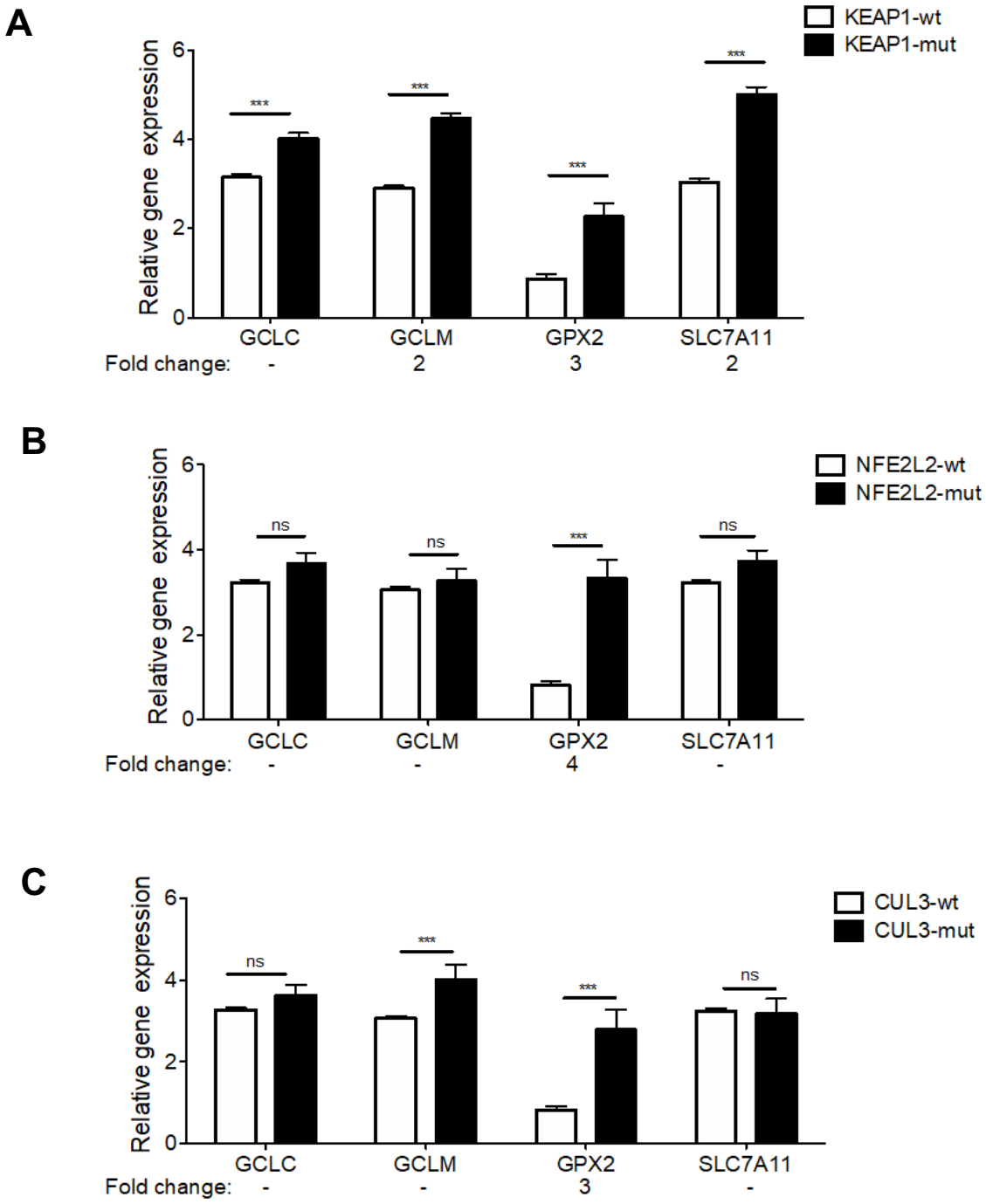


Figure 4.2.8: GSH-based antioxidant NRF2-target genes are more elevated in KEAP1-mut cell lines than cell lines harbouring mutations in NFE2L2 or CUL3

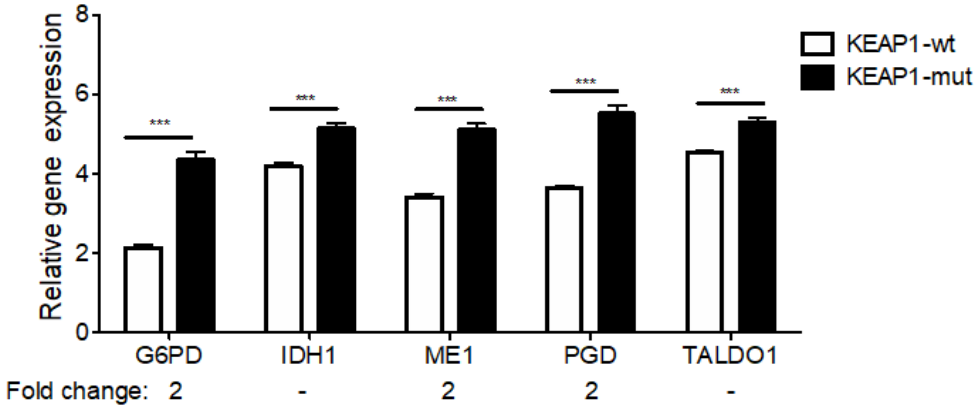
*The white bars on each of the graphs represents the cell line population that is wildtype (wt) for the gene in question and the black bars represents the cell line population that is mutant (mut) for the gene in question. Fold change values for each gene relative to the wt cell line group is displayed under each graph. (A) Analysis of the expression of four GSH-based antioxidant NRF2-target in cell lines that are either KEAP1-mut (28 cell lines) or KEAP1-wt (211 cell lines). The KEAP1-wt cell line population includes cell lines that harbor mutations in NFE2L2 and CUL3. No KEAP1-mut cell line also has a NFE2L2 mutation but one KEAP1-mut cell line has an additional CUL3 mutation. (B) Analysis of the expression of GSH-based antioxidant NRF2–target genes in cell lines that are either NFE2L2-mut (8 cell lines) or NFE2L2-wt (231 cell lines). The NFE2L2-wt cell line population includes cell lines that harbor mutations in KEAP1 and CUL3. One NFE2L2-mut cell line also harbors a mutation in CUL3. (C) Analysis of the expression of GSH-based antioxidant NRF2 –target genes in cell lines that are either CUL3-mut (8 cell lines) or CUL3-wt (231 cell lines). One CUL3-mut cell line also harbours a mutation in KEAP1 and a separate CUL3 mutant cell line has an additional mutation in NFE2L2. Bars on all three graphs represent the mean value obtained, and the error plotted is the associated SEM. All statistical analysis was carried out using GraphPad Prism 5 software. All significant increases are denoted with the * symbol; * represented a P< 0.05, ** P<0.01 and ** P< 0.001. P values > 0.05 were deemed not significant and are denoted by "ns".*

4.2.17 *KEAP1* mutant cell lines show significantly higher expression of NRF2-target genes associated with cell metabolism than cell lines harbouring mutations in *NFE2L2* or *CUL3*

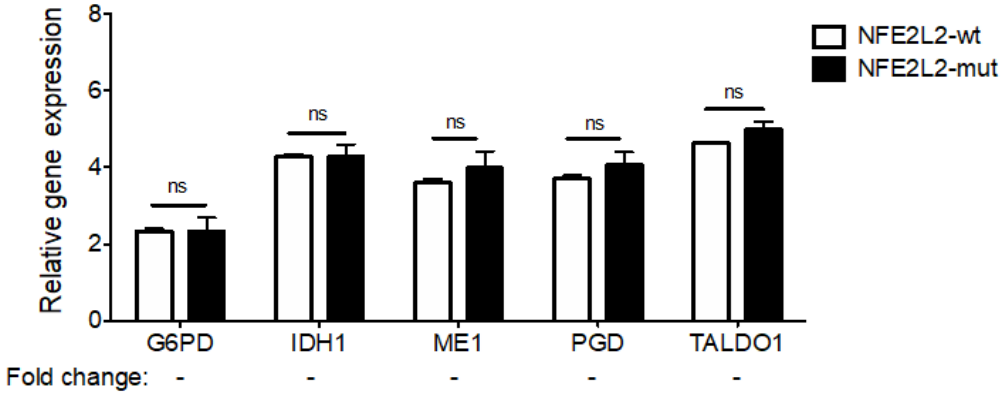
It has been reported that NRF2 influences cell metabolism through positively controlling the expression of genes involved in the pentose phosphate pathway (PPP) and in NADPH generation. To determine whether the presence of mutations in *KEAP1, NFE2L2* or *CUL3* affects expression of metabolism-related genes we analysed the expression of *G6PD, IDH1, ME1, PGD* and *TALDO1.* As shown in Figure 4.2.9A, *KEAP1*-mut cell lines have significantly higher expression of all six metabolism genes analysed in comparison to *KEAP1*-wt cell lines, with the largest fold change being seen in the expression of *G6PD*. However, *NFE2L2*-mut cell lines do not have significant alterations in the expression of any of metabolism related genes analysed when compared to *NFE2L2*-wt cell lines (Figure 4.2.9B). *CUL3*-mut cell lines show significantly higher expression of *ME1* and *TALDO1* but no significant alterations in the

expression of *G6PD, IDH1* or *PGD,* in comparison to *CUL3*-wt cell lines (Figure 4.2.9C).

A



B



C

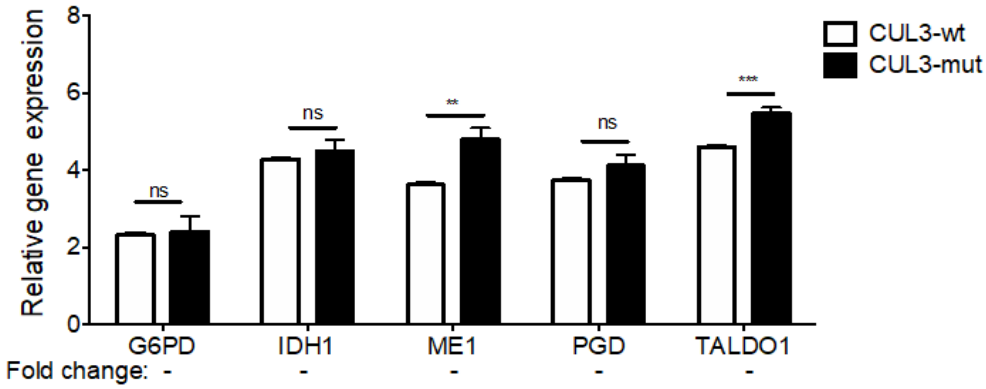


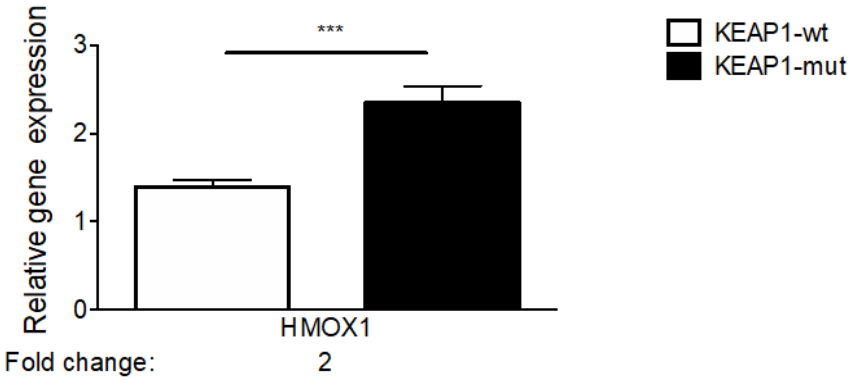
Figure 4.2.9: Genes associated with the PPP and NADPH generation are significantly elevated in cell lines with KEAP1 mutations in comparison to those with mutations in NFE2L2 or CUL3

*The white bars on each of the graphs represents the cell line population that is wildtype (wt) for the gene in question and the black bars represents the cell line population that is mutant (mut) for the gene in question. Fold change values for each gene relative to the wt cell line group is displayed under each graph. (A) Analysis of the expression of NRF2-target genes associated with the PPP and NADPH generation in cell lines that are either KEAP1-mut (28 cell lines) or KEAP1- wt (211 cell lines). The KEAP1-wt cell line population includes cell lines that harbor mutations in NFE2L2 and CUL3. No KEAP1-mut cell line also has a NFE2L2 mutation but one KEAP1-mut cell line has an additional CUL3 mutation. (B) Analysis of the expression of NRF2-target genes associated with the PPP and NADPH generation in cell lines that are either NFE2L2-mut (8 cell lines) or NFE2L2-wt (231 cell lines). The NFE2L2-wt cell line population includes cell lines that harbor mutations in KEAP1 and CUL3. One NFE2L2-mut cell line also harbors a mutation in CUL3. (C) Analysis of the expression of NRF2-target genes associated with the PPP and NADPH generation in cell lines that are either CUL3-mut (8 cell lines) or CUL3-wt (231 cell lines). One CUL3-mut cell line also harbours a mutation in KEAP1 and a separate CUL3 mutant cell line has an additional mutation in NFE2L2. Bars on all three graphs represent the mean value obtained, and the error plotted is the associated SEM. All statistical analysis was carried out using GraphPad Prism 5 software. All significant increases are denoted with the * symbol; * represented a P< 0.05, ** P<0.01 and ** P< 0.001. P values > 0.05 were deemed not significant and are denoted by "ns".*

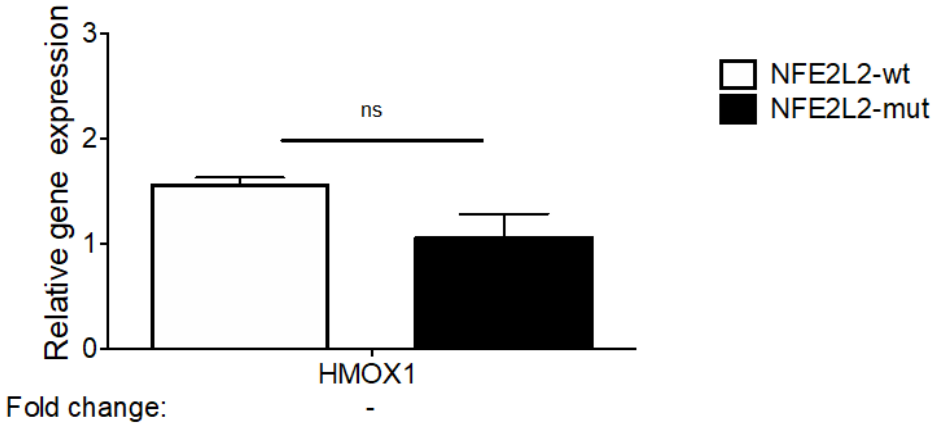
4.2.18 *KEAP1* mutant cell lines have significantly higher expression of the stress related NRF2-target gene *HMOX1*, in comparison to cell lines harbouring mutations in *NFE2L2* or *CUL3*

HMOX1 is a classical NRF2-target gene that is involved in stress pathways and iron metabolism, and has been implicated in the development of drug resistance. We analysed the expression of *HMOX1* in lung cancer cell lines harbouring mutations in either *KEAP1, NFE2L2* or *CUL3.* As shown in Figure 4.2.10, *KEAP*1-mut cell lines have significantly greater expression of *HMOX1* in comparison to *KEAP1*-wt cell lines, with a fold-change in expression of 2. Surprisingly, when we analysed the expression of *HMOX1* in cell lines with mutations in *NFE2L2* or *CUL3*, we found that neither of these mutations showed significantly higher expression of *HMOX1* in comparison to their corresponding wildtype cell line group.

A


***
KEAP1-wt
KEAP1-mut
Relative gene expression
3
2
1
0
HMOX1
Fold change:
2

B


ns
NFE2L2-wt
NFE2L2-mut
Relative gene expression
3
2
1
0
HMOX1
Fold change:
-

C

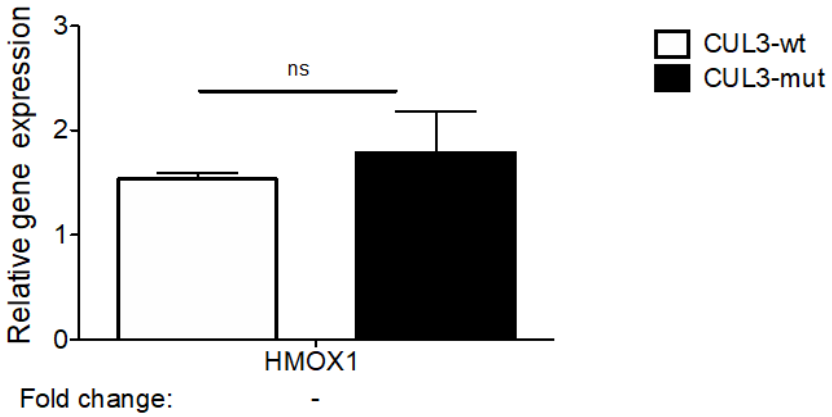

ns
CUL3-wt
CUL3-mut
Relative gene expression
3
2
1
0
HMOX1
Fold change:
-

Figure 4.2.10: Cell lines harbouring mutations in KEAP1 have significantly higher expression of the stress related NRF2-target gene, HMOX1, in comparison to cell lines harbouring mutations in NFE2L2 or CUL3

*The white bars on each of the graphs represents the cell line population that is wildtype (wt) for the gene in question and the black bars represents the cell line population that is mutant (mut) for the gene in question. Fold change values for each gene relative to the wt cell line group is displayed under each graph. (A) Analysis of the expression of HMOX1 in cell lines that are either KEAP1-mut (28 cell lines) or KEAP1-wt (211 cell lines). The KEAP1-wt cell line population includes cell lines that harbor mutations in NFE2L2 and CUL3. No KEAP1-mut cell line also has a NFE2L2 mutation but one KEAP1-mut cell line has an additional CUL3 mutation. (B) Analysis of the expression of HMOX1 in cell lines that are either NFE2L2-mut (8 cell lines) or NFE2L2-wt (231 cell lines). The NFE2L2-wt cell line population includes cell lines that harbor mutations in KEAP1 and CUL3. One NFE2L2-mut cell line also harbors a mutation in CUL3. (C) Analysis of the expression of HMOX1 in cell lines that are either CUL3-mut (8 cell lines) or CUL3-wt (231 cell lines). Bars on all three graphs represent the mean value obtained, and the error plotted is the associated SEM. All statistical analysis was carried out using GraphPad Prism 5 software. All significant increases are denoted with the * symbol; * represented a P< 0.05, ** P<0.01 and ** P< 0.001. P values > 0.05 were deemed not significant and are denoted by "ns".*

4.3 Discussion

Mutations in *KEAP1* and *NFE2L2* are often grouped together in terms of their effect on NRF2 signalling, and mutations in *CUL3* are often not studied. In this chapter we have carried out a comprehensive bioinformatics analysis and have made the following key findings: (1) *KEAP1* mutations are more prevalent than mutations in *NFE2L2* and *CUL3*; (2) *KEAP1* mutations co-occur with mutations *KRAS* and are mutually exclusive with mutations in *NFE2L2* and *TP53*; (3) mutations in *NFE2L2* co-occur with mutations in *TP53* and are mutually exclusive with mutations in *KEAP1* and *KRAS*; (4) mutations in *KEAP1* are predominantly homozygous missense mutations spread throughout the KEAP1 protein; (5) mutations in *NFE2L2* are predominantly heterozygous missense mutations that cluster to the DLG and ETGE motifs located in the Neh2 domain and (6) cell lines harbouring mutations in *KEAP1* have increased expression of NRF2-target genes in comparison to cell lines with *NFE2L2* or *CUL3* mutations.

4.3.1 Why mutations in *NFE2L2* and *KEAP1* do not occur together but you get *CUL3* mutations with *NFE2L2* mutations

We, like others, have found that mutations in *KEAP1* and *NFE2L2* are mutually exclusive to each other and never occur within the same cell line or tumour (The Cancer Genome Atlas Research, 2012; Shibata et al., 2008). It would be logical to presume that these mutations do not occur together because there is a level of redundancy with both mutations resulting in the same phenotype, namely, impaired degradation of NRF2 by KEAP1. Consistent with the previous findings of Fukutomi et al., (2014), we found mutations in *NFE2L2* predominately occur in the DLG and ETGE KEAP1-binding motifs that are located in the Neh2 domain of NRF2. There is strong evidence supporting the idea that mutations at these sites in NRF2 increase its activity through inhibiting KEAP1-mediated degradation of the transcription factor (Fukutomi et al., 2014).

Interestingly, we found that mutations in *NFE2L2* and *CUL3* do sometimes occur together in the same tumour, suggesting that the mutation in *CUL3* might potentially influence other cellular functions besides NRF2 signalling. CUL3 is one member of the cullin family of scaffold proteins that function recruit protein substrates for ubiquitination (Wimuttisuk et al., 2014). CUL3 binds to the BTB region of KEAP1 and KEAP1 binds NRF2. However, there are over 200 BTB domain containing proteins in human cells that all have the potential to be bound by CUL3 (Pintard et al., 2004). Therefore, some of the mutations in *CUL3* that we have analysed may not affect NRF2 degradation but the degradation of other oncogenic genes, meaning that mutations in *CUL3* may enhance oncogenesis independently of NRF2 and can occur alongside *NFE2L2* mutations.

4.3.2 *KRAS* mutations coexist with *KEAP1* mutations and *NFE2L2* mutations occur with *TP53* mutations

From the analysis described in this chapter we have discovered that mutations in *KEAP1* co-exist with mutations and *KRAS,* whereas mutations in *NFE2L2* tend to co-occur with mutations in *TP53*. Since mutations in *KEAP1* and *NFE2L2* mutations do not co-exist, there appears to be a segregation of each NRF2 influencing mutation with a different oncogenic driver mutation. In view of the fact that upregulation of NRF2 activity is not associated with cancer initiation but with tumour progression, it would be reasonable to suggest that the mutation

in *KRAS* or *TP53* arises as an early event that precedes mutation in *KEAP1* or mutation in *NFE2L2*, respectively.

KRAS is a member of the RAS family of evolutionally conserved GTPases. Mutations in *RAS* genes have been implicated in a wide range of cancers, with *KRAS* being the most commonly mutated family member. *KRAS* mutations are regularly associated with ADCs and are thought to arise through exposure to tobacco smoke (Ahrendt et al., 2001). Somatic mutations in *KRAS* commonly occur at either codon 12, 13 or 61 and result in the aberrant activation of KRAS signalling. Once activated KRAS signals to a wide range of downstream effector proteins that are involved in several signalling pathways that are implicated in cancer progression, such as the PI3K-AKT pathway that drives cell survival and the RAF-MEK-ERK pathway that is involved in cell proliferation (Westcott and To, 2012). Research has shown that oncogenic activating mutations in *KRAS* lead to the generation of high levels of mitochondrial ROS (Storz, 2016) through a variety of pathways such as activation of HIF-1α and HIF-2α signalling (Chun et al., 2010), suppression of respiratory chain complex I and III leading to a reduction in the mitochondrial transmembrane potential resulting in mitochondrial dysfunction (Hu et al., 2011b), and alterations in NADPH oxidase activity (Kong et al., 2013). However, when ROS increases to uncontrollable levels it triggers senescence or cell death pathways, so to ensure its survival the *KRAS* mutant cancer cell must suppress the heightened production of ROS so that it can benefit from the oncogenic influence of ROS without becoming subject to senescence or apoptosis. It has been suggested that the *KRAS* mutant cancer cells do this through upregulating NRF2 expression, by either directly increasing expression of the *NFE2L2* gene (DeNicola et al., 2011) or by acquiring somatic mutations in *KEAP1, NFE2L2* or *CUL3*. We like others, have shown that mutations in *KRAS* and *KEAP1* commonly occur together. Potentially this is because only a mutation in *KEAP1*, and not in *NFE2L2* or *CUL3*, is able to effectively combat the enhanced ROS produced when *KRAS* is mutated, but we currently have no data to support this.

P53 is a transcription factor that is activated upon DNA damage, which regulates the expression of a wide range of downstream target genes, of which many have contrasting functions. Several p53 target genes lead to cell cycle arrest allowing the cell time repair DNA damage but also expression of some p53

target genes has been shown to induce apoptosis of damaged cells in a ROS dependant manner. To permit the increase in ROS levels required for p53-induced apoptosis, p53 has been shown to suppress NRF2-dependant transcription of several ARE containing NRF2-target genes (Faraonio et al., 2007). Contrary to this idea, several p53-target genes, such as *SESN1* (also known as sestrin 1), *SESN2* (also known as sestrin 2), *SOD2* and *GPX1*, have been shown to decrease ROS levels. Interestingly, p53 has also been shown to regulate the expression of *GLS2* (also known as Glutaminase 2), which is the key enzyme involved in the conversion glutamine to glutamate, required for glutathione synthesis (Rotblat et al., 2012). It has also been shown that p53 may indirectly regulate NRF2 through stimulation of the NF-κB pathway (Tung et al., 2015) or directly through promoting the nuclear localization of the transcription factor (Lisek et al., 2018). By contrast, NRF2 has also been shown to regulate p53 expression. Murine double minute 2 (MDM2), the E3 ubiquitin ligase, that ubiquitinates p53 at several Lys residues leading to proteasomal degradation of the transcription factor, is responsible for maintaining low levels of p53 in resting cells (Gibbons et al., 2014). P53 itself can regulate the expression of MDM2 through a negative feedback loop but also NRF2 can control the expression of MDM2 through an ARE sequence (You et al., 2011), suggesting that active NRF2 will lead to increased levels of MDM2 activity and decreased expression of p53 protein. P53 is a tumour suppressor that becomes mutated in lung cancer, especially upon exposure to tobacco smoke (Gibbons et al., 2014), leading to oncogenesis typically characterized by uncontrolled cell division. This will lead to an accumulation of ROS, so to inhibit ROS induced apoptosis *TP53* mutant cancer cells increase NRF2 activity. We believe we are the first group to report that mutations in *TP53* show strong co-occurrence with mutations in *NFE2L2.* However, due to the complex feedback regulatory loop between NRF2 and p53 it is difficult to understand why a *NFE2L2* mutation would be selected instead of a mutation in *KEAP1.*

4.3.3 Why are somatic mutations in KEAP1 homozygous and those in NFE2L2 heterozygous?

One of the most interesting and exciting discoveries that has emerged from the research described in this chapter is that mutations in *KEAP1* are predominately homozygous whereas mutations in *NFE2L2* are predominantly

heterozygous. We will discuss this theme more in Chapter 5, but currently we have no data to explain why mutations in *KEAP1* and *NFE2L2* have different zygosities associated with them, but we believe this does in part account for their associated different levels of NRF2 activity. We speculate several potential reasons for this: (1) a homozygous mutation in *NFE2L2* have been evolutionally selected against; (2) mutations in *KEAP1* tend to be homozygous because KEAP1 functions as a dimer; (3) that to combat the high levels of ROS produced by *KRAS* mutations *KEAP1* mutations have to be homozygous whereas *TP53* mutations do not cause to such an elevation of ROS and only lead to homozygous mutations in *NFE2L2* and (4) mutations in tumour suppressors tend to be heterozygous indicating that *KEAP1* may be a tumour suppressor (this idea is discussed in more detail in the final discussion section, Chapter 6).

4.3.4 The idea that the frequency of *CUL3* and NFE2L2 mutations is low so we cannot make robust conclusions

In this chapter we have used a combination of three datasets in our cell line analysis, which has significantly strengthened our findings and enhanced the number of cell lines harbouring each mutation. However, this only gave us eight *NFE2L2* and eight *CUL3* mutant cell lines. These small numbers, means that we cannot be fully confident in the robustness of the data generated. To enhance our confidence, we have carried out the same analyses in tumour datasets. Using the tumour datasets not only increase the robustness of the data, because they have large n numbers, but also ensures our data are more clinically relevant. All tumour databases used in this chapter were generated by the TCGA consortium.

4.3.5 *KEAP1* mutant cell lines have very high drug detoxification gene expression

In Figure 4.2.7, we analysed the expression of the NRF2-target genes associated with drug detoxification, namely, *ABCB6, AKR1B10, AKR1C1, AKR1C2, AKR1C3* and *NQO1*, in cell lines with the presence of *KEAP1, NFE2L2* or *CUL3* mutations. Expression of all four AKRs (*AKR1B10, AKR1C1, AKR1C2* and *AKR1C3*) show significant elevation in expression between the *KEAP1*-mut and *KEAP1*-wt cell populations with large associated fold change values of 67, 5, 48 and 5 respectively. These were highest fold change values in expression for any gene analysed in the presence of any mutation. This suggests that *KEAP1* mutant cell lines selectively increased the expression of this subgroup of NRF2-

target genes, potentially this is due to them being highly resistant to platinum based drugs as the presence of *KEAP1* mutation has been demonstrated to dramatically influence platinum sensitivity (Tian et al., 2016). Also, overexpression of several members of the AKR family of genes has been documented in tumours from smokers. This is thought to be due to the AKRs being involved in the metabolism of the two major chemical carcinogens found in cigarette smoke (Penning, 2017). Since *KEAP1* mutations have been linked to smoking and occur alongside mutations in *KRAS* which are also linked with smoking, its maybe not surprising that we see increased expression of AKRs in *KEAP1* mutant cell lines. However, it should be highlighted that the expression data analysis was only done in lung cancer cell lines; which will have various smoking statuses, mixed histology and may or may not have been exposed to chemotherapy.

4.3.6 The limitations of this bioinformatics analysis

Bioinformatics is a useful biological tool to look for patterns in large data sets and identify potential new areas of research. However, there are several limitations to bioinformatics analysis, such as discussed previously the size of your dataset will influence the robustness of the data generated. Also, it is critical that all bioinformatics findings are validated experimentally.

We have analysed the expression of several NRF2-target genes in which we look at mRNA expression levels from several datasets that have been combined together, however these values do vary between the individual data sets. This could be due to the technique used to obtain and analyse the mRNA expression in each cell line and also by the cell culture conditions.

One limitation of our dataset, which will discuss again in Chapter 5, is that when we compare the expression of NRF2-target genes in the presence of a certain mutation we have not accounted for the presence of other mutations that will influence the expression of that gene. For example, in Figure 4.2.7 in the top image we compared the expression of six drug detoxification genes in the presence of a *KEAP1* mutation. In this analysis there are 28 cell lines that harbour a *KEAP1* mutation these cell lines do not harbour mutations in *NFE2L2* so all the *NFE2L2* mutant cell lines will be represented in the *KEAP1*-wt sub-group. The presence of a *NFE2L2* mutation would lead to the elevation in expression of all six genes analysed and hence the *KEAP1*-wt population will have enhanced

expression of these genes. So, we are in fact comparing the expression of NRF2 target genes between cell lines harbouring a mutation in *KEAP1* and cell lines that are wildtype for *KEAP1* mutation but harbour mutations in *NFE2L2.*

To correct for this, we could omit the *NFE2L2* mutant and the *CUL3* mutant cell lines from the *KEAP1* wildtype cell population and repeat the analysis and do the converse for the *NFE2L2* mutant analysis. Sadly, we would have had to do this filtering at the very beginning of our analysis while we still had access to the bioinformatics software. However, we have managed to reanalyse data looking at the expression of a subset of prototypic NRF2-target genes using the Sanger dataset this data will be presented in Chapter 5.

4.3.7 Additional data that was not included

Due to space constraints we have had to omit several pieces of bioinformatics data form this chapter. Including analysis of the expression of NRF2-target genes associated with: TXN-based antioxidants (*PRDX1, PRDX6, SRXN1* and *TXNRD1*), purine biosynthesis (*MTHFD2* and *PPAT*), iron metabolism (*FTH1* and *FLT*), proteasome subunits (*PSMA1, PSMA4, PSMB5* and PSMC1), serine/glycine biosynthesis (*PHGDH, PSAT1* and *SHMT2*), apoptosis (*BCL2*, also the non NRF2-target genes *BAD, BAX, BID* and *MCL1*) and autophagy (*ATG5, ATG7, CALCOCO2, SQSTM1* and *ULK1*). The expression these genes and the genes shown in this chapter, were also analysed in the presence of a *PTEN* or *PIK3CA* mutation.

4.3.8 Conclusions

Molecular profiling of the mutations implicated in lung carcinoma led to the discovery of the high frequency of *EGFR* mutations in ADC patients. Patients are now screened from the presence of an *EGFR* mutation and if deemed *EGFR* positive are treated with tyrosine kinase inhibitors (TKIs) with very promising results, highlighting the impact that genetic analysis can have on patient outcome. Since the discovery that aberrant expression of NRF2 leads to aggressive and chemo-resistant lung cancer, the field has predominantly focused on how to therapeutically decrease NRF2 levels and not on the genetic events that have led to the increased expression of the transcription factor. In this chapter through the use of bioinformatics dataset for lung cancer cell lines and tumours we have demonstrated the following: i) *KEAP1, NFE2L2* and *CUL3* mutations co-exist with different oncogenic mutations, mutations in these three genes map to regions of

each protein that will influence KEAP1 mediated degradation of NRF2; ii) *KEAP1*, *NFE2L2* and *CUL3* mutations are different in terms of their effect on the expression of NRF2-target genes; iii) *KEAP1* mutations are predominately homozygous whereas *NFE2L2* mutations are predominantly heterozygous. There is still a lot of work that needs to be carried out to validate these findings and to get a more comprehensive understanding of the impact that mutations in *KEAP1, NFE2L2* and *CUL3* play in lung carcinogenesis.

5. The identification of glutathione biosynthesis as a new selective therapeutic target in *KEAP1* mutant lung cancer cell lines.

5.1 Introduction

Glutathione is an abundant thiol-containing tripeptide found in almost all mammalian tissues that is composed of glutamic acid, cysteine and glycine (Meister and Anderson, 1983). Due to its potent antioxidant ability, glutathione has been implicated in the development and progression of several diseases, including cancer. The antioxidant activity of glutathione is due to the sulfhydryl group (SH) of the cysteinyl moiety functioning as a proton donor, giving glutathione its ability to control: the thiol status of other proteins, to detoxify electrophiles, drive cellular proliferation, and augment antioxidant defences (Gamcsik et al., 2012; Lu, 2009).

5.1.1 The role of glutathione in cancer

Glutathione levels have been documented to be altered in several cancer cell lines and tumours of various tissues of origins such as: liver, head and neck, ovary, colon, brain, breast, stomach, oesophagus and lung (Gamcsik et al., 2012). It is thought that the elevation in glutathione levels provides a pro-survival advantage to the cancer cell through inhibiting apoptosis, preventing senescence and conferring resistance to anticancer therapies.

5.1.1.1 Glutathione regulation of JNK and ASK-1 to control apoptosis

Anticancer therapies often exploit the toxic potential of oxidative stress. However, it should be noted that cancer cells produce a large amount of ROS due to their increased metabolic demand and the switch from oxidative phosphorylation to glycolytic metabolism (Vander Heiden et al., 2009; Nogueira and Hay, 2013). Both radiotherapy and chemotherapy lead to an increase in ROS production that is sufficient to cause cell death through apoptosis (Traverso et al., 2013). GST isoenzymes constitute a family of cytoprotective phase II drug-metabolizing enzymes that catalyse the addition of glutathione to either an anti-cancer drug or an endogenous protein through the formation of a thiol-ester bond (Tew, 1994). The GST family can be sub-divided into two families. These have been categorised as follows: the membrane-associated microsomal GST proteins (which are involved in leukotriene and prostaglandin metabolism,

sometimes called MAPEGs); the mitochondrial GST (sometimes called Kappa); and the cytosolic GST isoenzymes of which there are seven classes designated: Alpha, Mu, Pi, Sigma, Theta, Zeta and Omega (Hayes et al., 2005).

Amongst the cytosolic GST proteins, both class Mu and Pi enzymes have been implicated in the development of drug resistance through inhibiting MAPK signalling by non-enzymatic protein-protein interactions (Townsend and Tew, 2003). JNK and ASK-1 are both inhibited under homeostatic conditions and once activated lead to the stimulation of signalling cascades resulting in apoptosis. GST P1-1 is the most abundant of all the transferases and has been reported to be overexpressed in several cell lines that have acquired resistance *in vitro* through repetitive exposure to antineoplastic drugs. JNK activity is low under non-stressed conditions through a C-terminal association with GST P1-1. The converse of this has also been demonstrated with exposure to ultraviolet (UV) light leading to activation of JNK through inhibition of GST P1-1 (Adler et al., 1999; Wang et al., 2001). ASK-1 has been shown to be activated under conditions of heat shock through the dissociation of ASK-1 from GST M1-1 (Dorion et al., 2002). Thioredoxin has also been shown to regulate ASK-1 expression, similar to GST M1. However, thioredoxin mediated regulation of ASK-1 is redox-dependent (Saitoh et al., 1998).

Rapid cellular proliferation is one of the classical hallmarks of cancer. Studies carried out in both cell culture models and patients have revealed that treatment with the glutathione biosynthesis inhibitor BSO reduces cell proliferation through stimulation of apoptotic pathways, such as the JNK pathway (Schnelldorfer et al., 2000). This has been confirmed through the use of GST P1/P2 KO mice which show a reduced proliferative response, due to aberrant JNK signaling (Ruscoe et al., 2001). Also glutathione has been shown to regulate cell proliferation through apoptosis-independent pathways, potentially through alterations in NF-κB signaling via IKKβ (Reddy et al., 2007).

5.1.1.2 The role of glutathione in drug resistance

Over-expression of several GST isoenzymes in mammalian cells leads to the development of drug resistance (Puchalski and Fahl, 1990). SCLC cell lines and tumours have low innate resistance to antineoplastic drugs but often develop multi-drug resistance. Examination of such cells led to the discovery of the ATP-dependant ABC transporter, MRP, which was shown to be overexpressed in a

SCLC cell line that had acquired resistance to doxorubicin (Cole et al., 1992). MRP has now been shown to be responsible for the transport of several electrophilic anti-cancer drugs (especially alkylating agents) conjugated to glutathione, such as cisplatin (Lan et al., 2018). This conjugation of glutathione with electrophilic drugs serves two important functions: firstly, it inactivates the toxic properties of the drug, secondly, it makes the drug a substrate for the pump (Gamcsik et al., 2012). MRP1, the most well studied and characterised of the MRP family of transporters, is often overexpressed in oncogenic tissues and is thought to lead to resistance through enhancing export of anti-cancer drugs out of the tumour cell (Cole, 2014; Cole and Deeley, 2006).

MRP has also been shown to be involved in the transport of both reduced and oxidized glutathione, and hence is involved in cellular health through the maintenance of the reduced to oxidized glutathione ratio and ensuring high intracellular concentrations of reduced glutathione (Rappa et al., 1998).

5.1.2 The glutathione biosynthetic pathway

The levels of total glutathione in the cell are tightly regulated by balancing its production and utilization. The *de novo* biosynthesis of glutathione is a two-step ATP-dependant process (Figure 5.1.1). The first rate-limiting step involves the formation of γ-glutamylcysteine from glutamic acid (which is produced by the enzymatic conversion of glutamine by the enzyme glutaminase (GLS)) and cysteine. This is catalysed by glutamate-cysteine ligase (GCL) which is a heterodimeric enzyme composed of a catalytic subunit, the glutamate-cysteine ligase catalytic (GCLC) subunit and a modifier subunit, the glutamate-cysteine ligase modifier (GCLM) subunit. The second stage involves production of glutathione from γ-glutamylcysteine and glycine, catalysed by glutathione synthase (GS). The production of glutathione in the cell under homeostatic conditions is dependent on the availability of dietary cysteine and GCL activity (Traverso et al., 2013;Daseul Kim et al., 2015). Once transported into the cell through the transporter xCT, cystine rapidly auto-oxidizes to cysteine, leading to the production of ROS. Therefore, the majority of the cysteine in the cell originates from the γ-glutamyl cycle, in which γ-glutamyltranspeptidase (GGT) breaks down the peptide bond in releasing cysteine (Lu, 2009; Meister, 1983). Cysteine can also be produced from methionine in a process called the cystathionine pathway (Meister and Anderson, 1983). However, this pathway is

only found to be active in hepatic cells and is redundant in all other cell types (Tarver and Schmtdt, 1939).

Glutathione is found in either one of two states in the cell: reduced glutathione (GSH) or oxidized glutathione (GSSG) (Kaplowitz et al., 1985) (Figure 5.1.2). The majority of glutathione in the cell is in the reduced GSH form which reaches intracellular concentrations in the milli-molar range. By contrast, the levels of GSSG in the cell represent approx. only 1% of the total glutathione and only reach the micro-molar range (Akerboom et al., 1982).

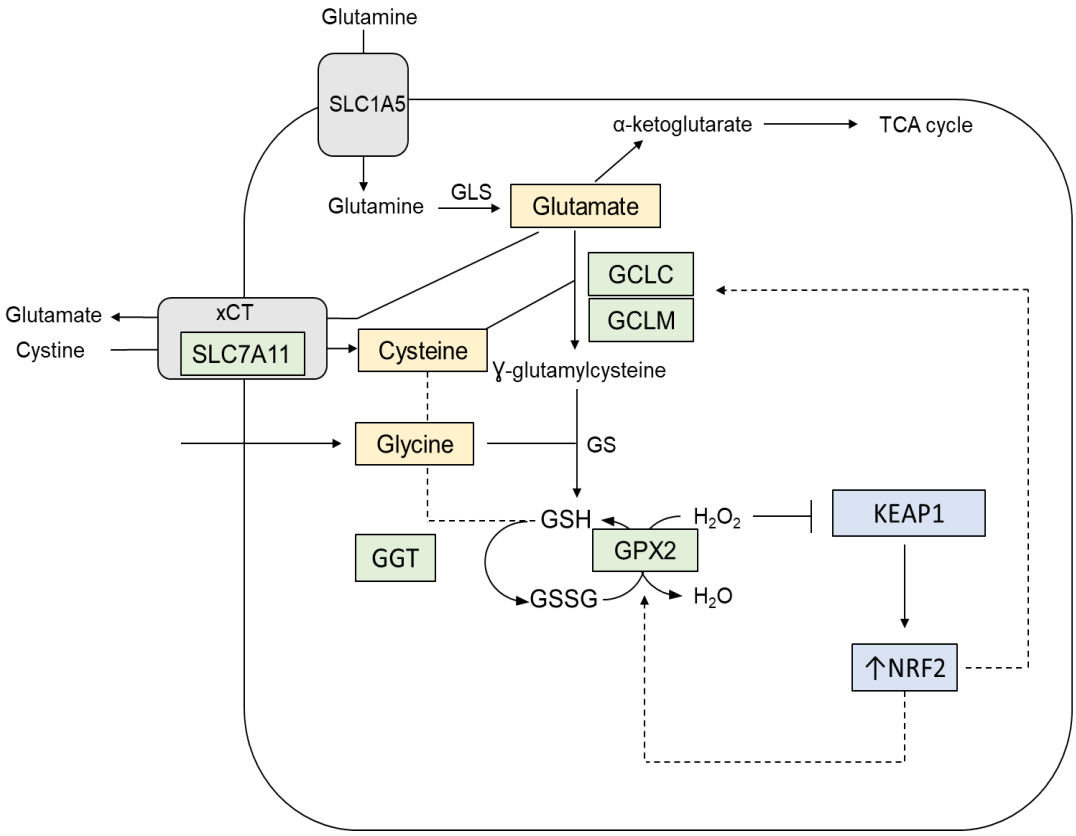


Figure 5.1.1: Schematic diagram showing the influence that NRF2 plays on the production of glutathione in the cell

Glutamine is transported into the cell through the transporter SLC1A5 and rapidly converted to glutamate, catalysed by GLS. Glutamate can then be either; converted to α-ketoglutarate, be exported out of the cell in exchange for the import of cystine through the transporter xC^- or used to produce glutathione. To produce glutathione, glutamate is converted to γ-glutamylcysteinine by GCL. γ-glutamylcysteine and glycine are then converted to glutathione in a reaction catalysed by GS. GPX2 uses GSH as a co-factor for the neutralisation of H_2O_2. The pool of cysteine in the cell can be replenished through GGT breakdown of GSH. The three building blocks of glutathione are highlighted in yellow boxes. Highlighted in the green coloured boxes are NRF2-target genes; GCLC, GCLM, SLC7A11, GPX2 and GGT, which all play fundamental roles in the production and utilization of the glutathione in the cell.

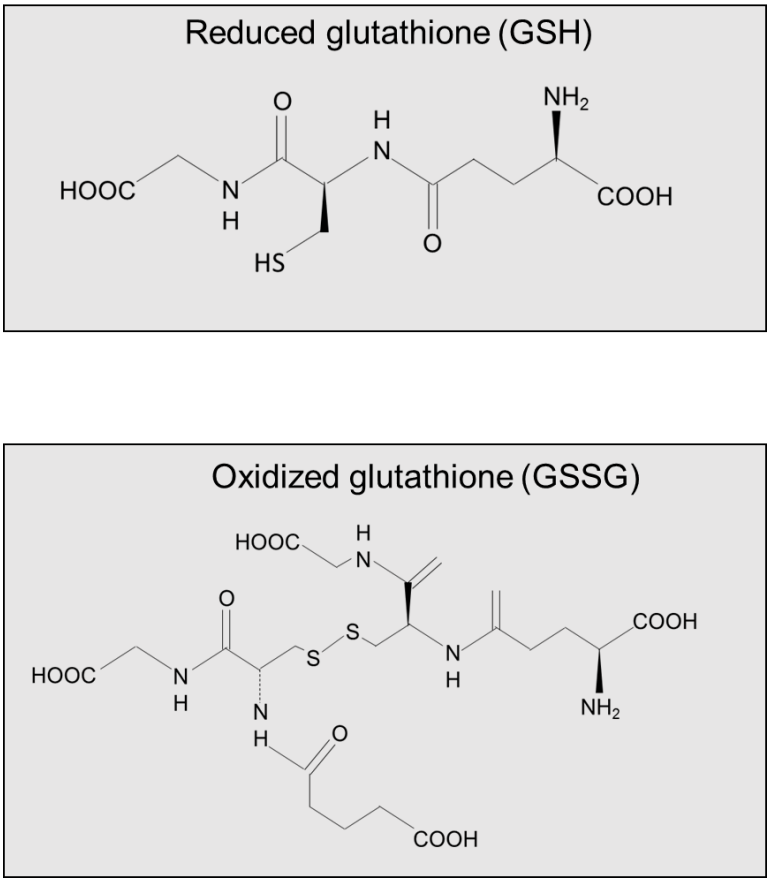


Figure 5.1.2: The chemical conversion of reduced glutathione to oxidized glutathione

GSH shown in the top grey box can be converted to oxidized glutathione shown in the bottom grey box. GSSG is composed of two molecules of GSH linked by a disulphide bond. Once oxidized, the reaction can occur in the reverse catalysed by glutathione reductase (GR) using NADPH as a co-factor, producing GSH. The ratio of GSH to GSSG is often used as an indicator of oxidative stress.

5.1.3 The role of NRF2 in regulating glutathione levels

The transcription factor NRF2 functions as a master regulator of the oxidative stress defences of the cell through its ability to transcriptionally activate of over 200 cytoprotective genes, including those involved in both the production and utilization of glutathione (Tebay et al., 2015). All NRF2-target genes possess at least one ARE-sequence in their promoter regions (Rushmore and Pickett, 1990; Rushmore et al., 1991). Pioneering work by Masi Yamamoto and colleagues demonstrated the influence that NRF2 plays on the expression of ARE-containing genes through the generation of a knockout mouse (Itoh et al., 1997). They provided evidence that GST gene induction by butylated BHA was diminished in the absence of NRF2, suggesting that the transcription factor regulates expression of multiple drug detoxification genes. In subsequent years, ARE sequences have been discovered in multiple important detoxification genes such as *GSTA1, GSTM1, GCLC, GCLM* and *NQO1*, which have also been shown to be transcriptionally regulated by NRF2.

5.1.3.1 NRF2 in glutathione production

The initial rate limiting step in *de novo* biosynthesis of glutathione is catalysed by the enzyme GCL, which is composed of two subunits encoded by two separate genes *GCLC* and *GCLM*. GCLC, is a 73 kDa subunit that is responsible for the catalytic activity of GCL and is also involved in the feedback inhibitory loop controlling the levels of glutathione produced (Lu, 2013). GCLM is a 28 kDa subunit that lacks catalytic activity but regulates the kinetic properties of GCLC. Both the *GCLC* and *GCLM* genes contain ARE sequences in their promoter regions and not surprisingly are therefore upregulated in the presence of oxidative stress (Erickson et al., 2002). Kenny MacLeod and colleagues carried out a siRNA knock down of *KEAP1* in human keratinocyte HaCaT cells followed by a microarray to identify candidate NRF2-target genes. In total they identified 23 genes that were increased over 2-fold in comparison to the control, which included both *GCLC* and *GCLM*. Under cell culture conditions, the intracellular level of glutathione was found to be increased in HaCat cells when *KEAP1* was silenced and decreased when NRF2 was knocked out, suggesting that NRF2 regulates glutathione levels *ex vivo* through controlling expression of GCLC and GCLM (MacLeod et al., 2009). Using knockout mice, others have reported that Nrf2 controls the expression of both GCLC and GCLM subunits of GCL

(Thimmulappa et al., 2002) and loss of Nrf2 has also been shown to result in low intracellular levels of glutathione (Morales-Pantoja et al., 2016; Blake et al., 2010; Chan et al., 2001).

The availability of cysteine is a limiting factor in glutathione biosynthesis. xC^- is a cell surface antiporter, that transports extracellular cystine into the cell coupled to the anti-port of intracellular glutamate at a ratio of 1:1(Koppula et al., 2018). The xC^- transport system is a heterodimer composed of two proteins; solute carrier family 7 member 11 (xCT, encoded by the gene *SLC7A11*) and solute carrier family 3 member 2 (encoded by the gene *SLC3A2*). SLC7A11 is responsible for the transporter ability of the xC^- complex, whereas SLC3A2 is responsible for the recruitment of SLC7A11 to the plasma membrane (Koppula et al., 2017). Work by Shiro Bannai and colleagues, identified several ARE-sequences in the promoter region of *SLC7A11* and demonstrated that NRF2 can regulate the expression of this gene (Sasaki et al., 2002). This has been further validated in breast cancer cell lines, demonstrating that under conditions in intracellular oxidative stress xC^- is upregulated due to the nuclear translocation of NRF2 leading to the active transcription of *SLC7A11* (Verschoor and Singh, 2013; Habib et al., 2015).

5.1.3.2 NRF2 in glutathione utilization

The total levels of cellular glutathione are not only being dictated by its production but also by its consumption. GPX isoenzymes are selenium–dependant enzymes that protect membranes from ROS through catalysing the reduction of H_2O_2, or organic hydroperoxides, to water and alcohols by utilizing GSH as a co-factor. There are currently eight human GPX isoenzymes, known as GPX1-8, which all show different localizations and substrate specificity (Margis et al., 2008). GPX2, also known as gastrointestinal GPX, has been shown to contain a functional ARE-sequence and its expression is reduced in Nrf2 KO mice (Banning et al., 2005). Also GPX2 has been revealed to be induced in response to carcinogenic cigarette smoke, through cigarette smoke inhibiting the interaction between KEAP1 and NRF2 leading nuclear accumulation of NRF2 and the active transcription of *GPX2* (Singh et al., 2007) .

5.1.4 Alterations in NRF2 signalling in lung cancer

NRF2 is often elevated in several types of cancer and this is associated with poor prognosis (Shibata et al., 2008). Lung cancer is a classic example of a

cancer where increased NRF2 signalling confers an advantage to the cancer cell through increasing cellular proliferation, metabolism and the ability to withstand oxidative stress. This upregulation of the transcription factor is commonly due to somatic mutations in the gene encoding the negative regulatory protein KEAP1 or mutations in *NFE2L2* (mutations that lead to increased NRF2 activity are discussed in detail in Chapter 4). Mutations in both of these genes are often grouped together in their ability to halt KEAP1-dependent proteasomal degradation of NRF2, and thus stimulate NRF2-dictated gene expression (Hast et al., 2014). Expression of genes associated with glutathione biosynthesis and utilization, namely *GCLC, GCLM, SLC7A11* and *GPX2,* have been shown to be elevated in lung cancer in the presence of the aforementioned mutations in *KEAP1* and *NFE2L2*.

To date, work in the field has clearly linked an increase in NRF2 activity found in lung cancer, due to mutations in *KEAP1* and *NFE2L2*, to an increase in intracellular glutathione levels through increased expression of the NRF2-target genes such as *GCLC, GCLM* and *SLC7A11*. However, to date, no one has looked at the consequences of *KEAP1* and *NFE2L2* mutations separately to determine whether they exert similar effects in terms of their effect on glutathione. The data in this chapter characterises several differences between *KEAP1* mutant and *NFE2L2* mutant cell lines and highlights the selective therapeutic opportunity of targeting the glutathione pathway in *KEAP1* mutant cell lines when compared with *NFE2L2* mutant cell lines.

5.2 Results

5.2.1 Several of the most commonly mutated genes in lung cancer encode proteins that have the potential to activate NRF2

Cancer does not originate from one mutation in a single gene but through the combination of several mutations in range of different genes for proteins with critical functions. High levels of NRF2 activity are found in both lung cancer cell lines and tumours, and is associated with poor prognosis (Shibata et al., 2008). This upregulation in NRF2 activity can arise through several different routes but one of the most common is mutations in the genes encoding KEAP1 and NRF2 (Tebay et al., 2015). Although, these are not the only mutations that will result in an elevation in NRF2 signalling. As shown in Figure 5.2.1, mutations in genes

encoding the proteins EGFR, STK11 (gene name *LKB1*), PIK3CA, and PTEN are also likely to lead to an elevation in the abundance and activity of the transcription factor through inhibiting $SCF^{\beta\text{-}TrCP}$-mediated degradation. Also, mutations in *CUL3*, which are often not included in the literature, will lead to an increase in NRF2 levels through impairing KEAP1-dependent degradation.

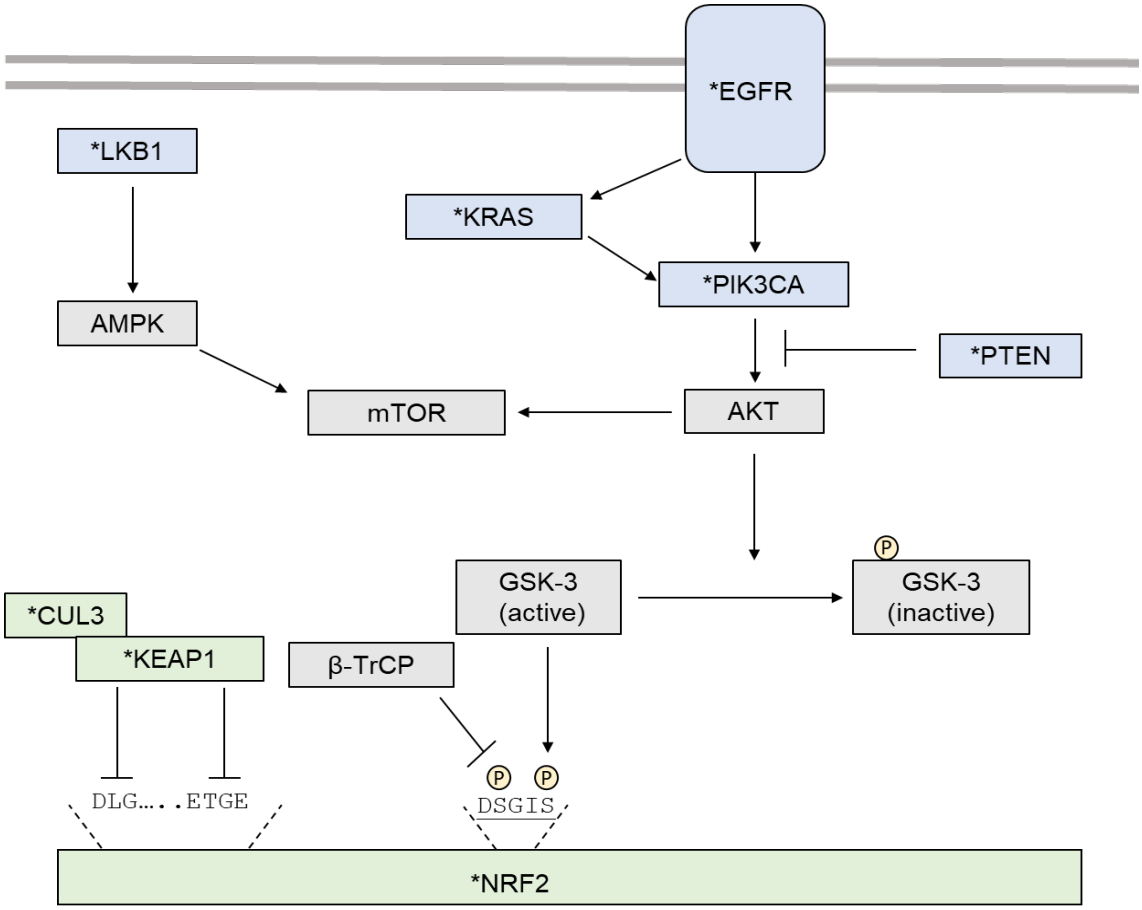


Figure 5.2.1: Several proteins whose genes are commonly mutated in lung cancer activate the NRF2-signaling pathway

*Pathway diagram showing the interplay between proteins whose genes are commonly mutated in lung cancer and how they link to NRF2-signalling. All proteins donated with a * are encoded by genes that are commonly mutated in lung cancer. The light blue cascade on the far right-hand side highlights how mutations in; EGFR, KRAS, PIK3CA, PTEN and STK11 (the gene that encodes the protein LKB1) influence the degradation of NRF2 through altering GSK-3 phosphorylation and inhibiting the $SCF^{\beta\text{-}TrCP}$ axis. The green boxes on the left-hand side of the diagram represents how mutations in CUL3 and KEAP1 halt 26S proteasomal degradation of NRF2. Mutations can also occur in NFE2L2 which leads to changes in the DLG and ETGE motifs, and flanking amino acids, that interact with KEAP1.*

5.2.2 In lung cancer, somatic mutations in *KEAP1* are more prevalent than those in *NFE2L2* and have a greater effect on NRF2-target gene expression.

To determine whether the presence of a *KEAP1* mutation has the same effect as a *NFE2L2* mutation, in terms of NRF2 signalling, we carried out a bioinformatics analysis of lung cancer cell lines using NRF2-target gene expression as a read out of NRF2-signalling. Using the software TIBCO Spotfire® we combined Sanger, AstraZeneca, and CCLE datasets. Focusing on pan lung cancer (i.e., no specific histological subtype) we filtered the cell lines for the presence of mutations in *KEAP1* and *NFE2L2*. We found that none of the cell lines in the datasets harboured both a mutation in *KEAP1* and *NFE2L2*. Out of 239 cell lines in the datasets, 28 harbour mutations in *KEAP1* whereas only 8 harbour mutations in *NFE2L2*, making *KEAP1* mutations over three times more prevalent than *NFE2L2* mutations (Table 5.2.1).

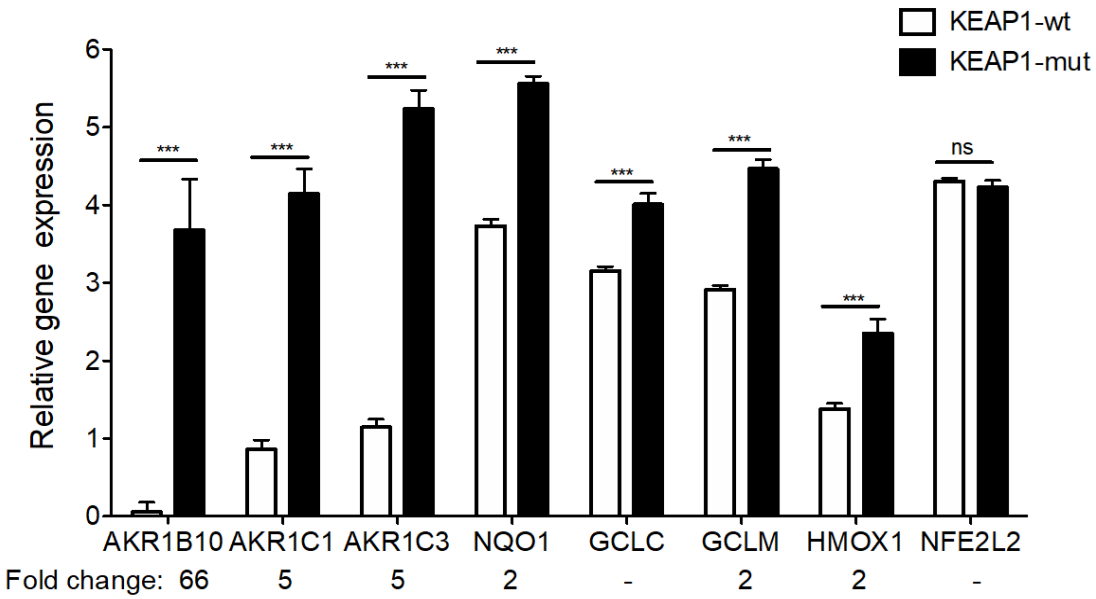
We then compared the expression of seven prototypic NRF2-target genes, namely, *AKR1B10, AKR1C1, AKR1C3, NQO1, GCLC, GCLM* and *HMOX1*, between cell lines that were grouped together as follows: those harbouring mutations in *KEAP1* (i.e., *KEAP1*-mut), wildtype (wt) for *KEAP1* (i.e., *KEAP1*-wt), harbouring mutations in *NFE2L2* (i.e., *NFE2L2*-mut) or wt for *NFE2L2* (i.e., *NFE2L2*-wt). As shown in Figure 5.2.2A and 5.2.2B, the *KEAP1*-mut cell lines show a greater elevation in expression of all the seven NRF2-target genes analysed when compared with the *NFE2L2*-mut cell lines. One of the most pronounced changes between the *KEAP*1-mut and *NFE2L2*-mut cell lines is that of the expression of *AKR1B10*. Neither *KEAP1*-mut nor *NFE2L2*-mut cell lines show an elevation in mRNA for *NFE2L2* in comparison to their wildtype counterparts. This is to be expected as changes in NRF2 signalling will only be affected at a translational level and not a transcriptional level.

Name of gene	Number of cell lines harboring mutations (Total number of cell lines 239)							
	KEAP1	*KRAS*	*NFE2L2*	*CUL3*	*PIK3CA*	*PTEN*	*STK11*	*EGFR*
KEAP1	**28**	12	0	1	3	0	5	2
NFE2L2	0	0	**8**	1	1	0	0	0

Table 5.2.1: Frequency of mutation in either KEAP1 or NFE2L2 amongst the lung cancer cell line dataset

Bold font indicates the total number of KEAP1(28) and NFE2L2 (8) mutant cell lines out of the 239 analysed. The rest of the table indicates the number of instances that KEAP1 and NFE2L2 mutations with mutations co-occur with mutations in other genes that may alter NRF2 degradation.

A



B

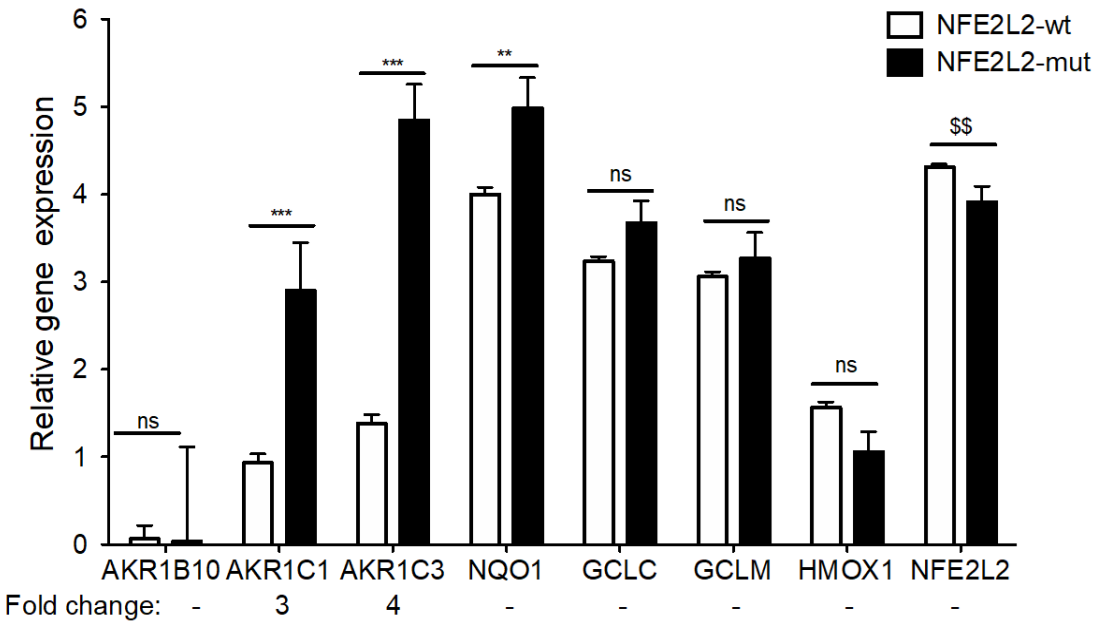


Figure 5.2.2: Somatic mutations in KEAP1 are more prevalent and have a greater effect on NRF2-target gene expression than those in NFE2L2.

*The black bars on each graph represent the mutant cell line population and the white bars represent the wildtype population. The names of all seven genes analysed is given on the x-axis. The fold-change in expression of particular ARE-driven genes within the cell lines harbouring mutant KEAP1 or NFE2L2 relative to the expression in the remaining cell lines harbouring wildtype KEAP1 or NFE2L2 is displayed underneath the graph A horizontal dash represents no significant fold-change in gene expression. Significant changes in gene expression in the mut cell lines relative to the wt cell lines are displayed on each graph. (A) Analysis of the expression of seven prototypic NRF2-target genes in lung cancer cell lines that are either KEAP1-mut (28 cell lines) or KEAP1-wt (211 cell lines) (B) Analysis of the expression of seven prototypic NRF2-target genes in lung cancer cell lines that are either NFE2L2-mut (8 cell lines) or NFE2L2-wt (231 cell lines). Data presented on both graphs represent the mean value obtained, and the error plotted is the associated SEM. Statistical analysis was carried out using GraphPad Prism 5 software. All significant increases are denoted as follows: *, P< 0.05; **, P<0.01; *** P< 0.001. P values > 0.05 were deemed not significant and are denoted by "ns". Significant decreases are denoted with a $ sign.*

5.2.3 Lung cancer cell lines in the Sanger data set that harbour mutations in *KEAP1* display higher levels of expression of NRF2-target genes than those harbouring *NFE2L2* mutations.

As no one cell line poses both a mutation in *KEAP1* and *NFE2L2*, this means that in the previous analysis, shown in Table 5.2.1, *KEAP1*-mut cell lines will be compared against cell lines that are wt for *KEAP1* mutation but also some cell lines will harbour *NFE2L2* mutations. This may not have a pronounced effect on the data generated for the *KEAP1* analysis because there are only eight cell lines that harbour a *NFE2L2* mutation, meaning that out of the 211 *KEAP1*-wt cell lines, 203 will be wildtype for both *NFE2L2* and *KEAP1* mutation. However, this could have a large impact on the analysis of *NFE2L2* mutations, because when we compare the expression of ARE-driven genes in eight *NFE2L2*-mut cell lines with expression in the 231 *NFE2L2-wt* cell lines*,* 28 of *NFE2L2*-wt cell lines will possess a mutation in *KEAP1* and only 203 will be wildtype for both *KEAP1* and *NFE2L2* mutation. This means the changes in gene expression may not look as pronounced because the *NFE2L2*-wt population will have elevated expression because it includes cell lines with mutant *KEAP1.*

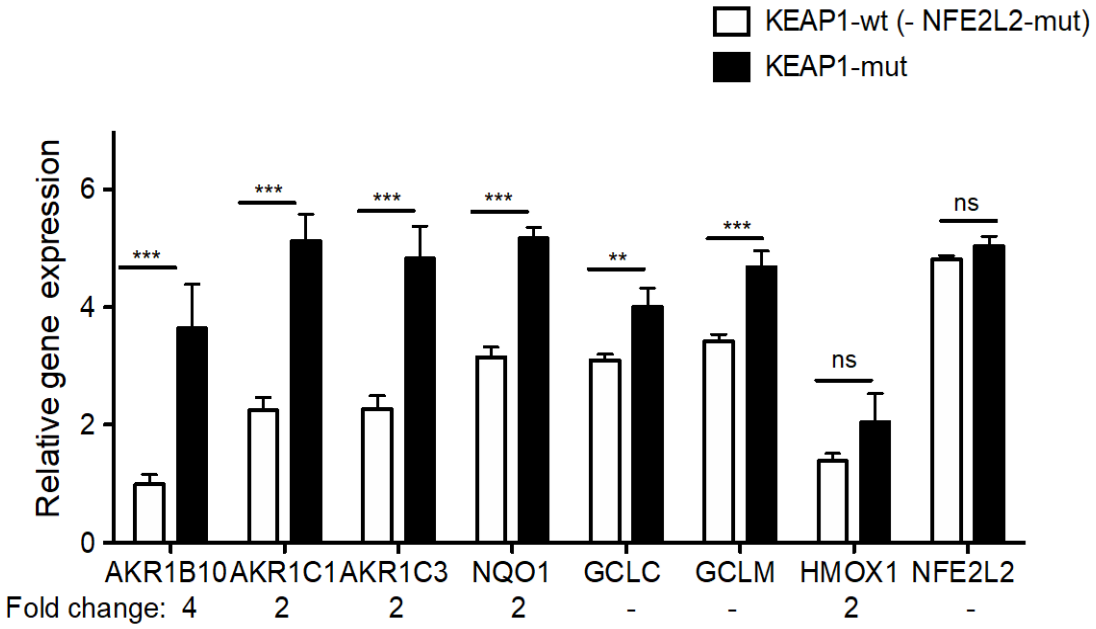
To try to account for this 'contamination' of the wild-type dataset we reanalysed expression of our chosen seven prototypic NRF2-target genes using only the Sanger dataset. Table 5.2.2 highlights the major caveat of this analysis, as due to analysing only one dataset we were not able to get all the data points we had in the first analysis because the number of cell lines is limited. We identified 14 cell lines in the Sanger dataset that harbour mutations in *KEAP1* and four that harbour mutations in *NFE2L2* out of the 122 cell lines in total. As shown in Figure 5.2.3A and B, the overall the magnitude of NRF2-target gene expression across the two mutant populations appears to be less than observed in Figure 5.2.2A and B. It was found that *KEAP1* mutant cell lines still had significantly higher expression of six out of the seven NRF2-target genes in comparison to the *KEAP1/NFE2L2*-wt cell lines (i.e. *KEAP1*-wt lines from which those with mutant *NFE2L2* had been removed). The only gene to be changed by this analysis, when compared with Figure 5.2.2A, is *HMOX1* which is not significantly altered between *KEAP*1-mut and *KEAP1/NFE2L2*-wt cell groups. However, the elevation in NRF2-target gene expression in the *NFE2L2*-mut cell lines does appear to be dramatically reduced when the *KEAP1*-mut cell lines were removed from the

analysis. Nevertheless, the overall pattern still holds with *KEAP1*-mut cell lines showing enhanced NRF2-target gene expression relative to the *NFE2L2*-mut cell lines.

Number of cell lines harbouring mutations			
KEAP1-mut	*KEAP1*-wt (-*NFE2L2*-mut)	*NFE2L2*-mut	*NFE2L2*-wt (-*KEAP1*-mut)
14	108	4	94

Table 5.2.2: Indicating the number of cells lines that harbor a mutation in KEAP1 and NFE2L2 out of the 122 cell lines analyzed using the Sanger dataset

A



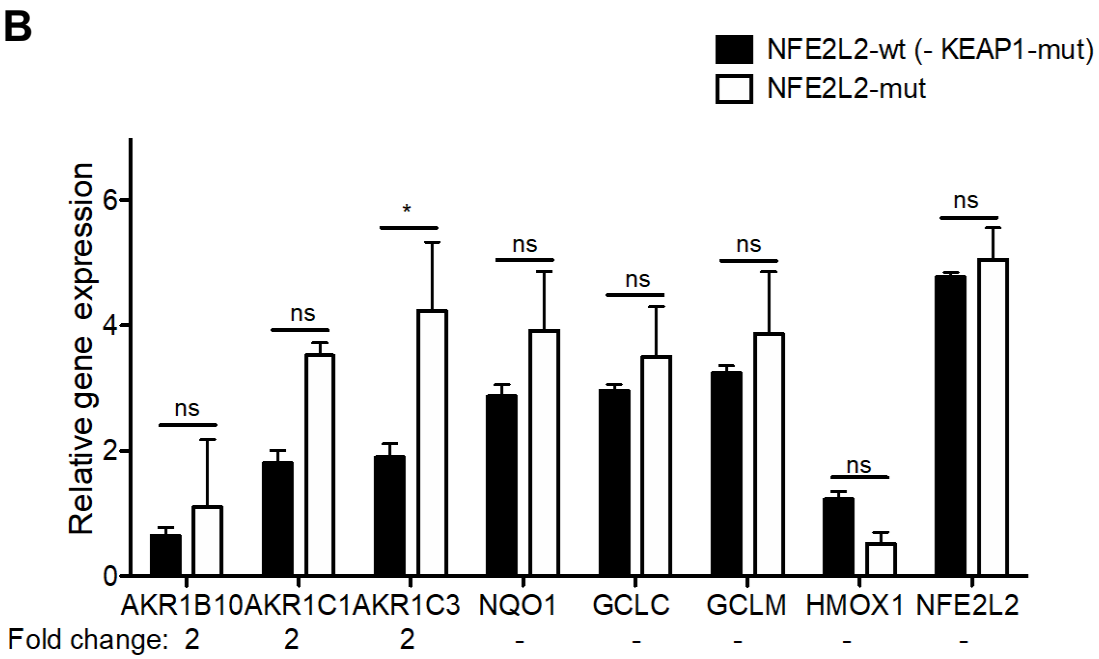


Figure 5.2.3: Lung cancer cell lines in the Sanger dataset that harbour mutant KEAP1 show higher expression of NRF2-target genes than those with mutations in NFE2L2

*The black bars on both graphs represent the mutant cell line population and the white bars the wildtype population. The names of all seven genes analysed is given on the x-axis. The fold-change in expression of particular ARE-driven genes within the cell lines harbouring mutant KEAP1 or NFE2L2 relative to the expression in the remaining cell lines harbouring wildtype KEAP1 or NFE2L2 is displayed underneath the graph. A horizontal dash indicates no significant fold-change in gene expression. Significant changes in gene expression in the mut cell lines relative to the wt cell lines are displayed on each graph. (A) Analysis of the expression of seven prototypic NRF2-target genes in lung cancer cell lines that are either KEAP1-mut (14 cell lines) or KEAP1-wt (-NFE2L2-mut) (108 cell lines). (B) Analysis of the expression of seven prototypic NRF2-target genes in lung cancer cell lines that are either NFE2L2-mut (4 cell lines) or NFE2L2-wt (- KEAP1-mut) (94 cell lines). Data presented on both graphs represent the mean value obtained, and the error plotted is the associated SEM. Statistical analysis was carried out using GraphPad Prism 5 software. Significant increases are denoted as follows: *, P< 0.05; **, P<0.01; *** P< 0.001. P values > 0.05 were deemed not significant and are denoted by "ns". Significant decreases are denoted with a $ sign.*

5.2.4 The diverse genetic composition of the six commercially-available cell lines used to validate the bioinformatics analyses.

To validate the findings made in the bioinformatics analysis, six commercial cells lines were analysed in terms of their NRF2- signalling. As shown in Table 5.2.3, the six cell lines included two *KEAP1*-mut cell lines, two *NFE2L2*-mut cell lines and two "wildtype" cell lines with mutations in neither *KEAP1* nor *NFE2L2.* These were as follows: representing *KEAP1*-mut cell lines, A549 and H460 cell lines both harbour homozygous mutations in *KEAP1* and homozygous mutations in *KRAS* (also H460 cells also have a heterozygous mutation in *CUL3* and a heterozygous mutation in *PIK3CA*); representing *NFE2L2*-mut cell lines, H2228 and H1568 cell lines possess heterozygous mutations in *NFE2L2* and homozygous mutations in *TP53*; representing "wildtype" cell lines, H1299 and H1395 are wildtype for genes that are known to effect NRF2-signalling (notably, H1299 cell lines cells have a partial deletion in the gene for p53). It should be noted at this time that the function of some of these mutations is unknown, but all are thought to be oncogenic.

Table 5.2.4, demonstrates the histological diversity of the six cell lines chosen. The majority of the cell lines are derived from adenocarcinomas (A549, H2228, H1568, H1395), though H460 is a large cell carcinoma cell line, and H1299 is from a lymphoid metastasis.

Cell line	Name of mutant genes, the location of mutations and whether Homozygous / Heterozygous							
	KEAP1	*NFE2L2*	*CUL3*	*PTEN*	*PIK3CA*	*STK11*	*KRAS*	*EGFR*	*TP53*
A549	p.G333C (Homozygous)	WT	WT	WT	WT	WT	p.G12S (Homozygous)	WT	WT
(NCI)H460	p.D236H (Homozygous)	WT	p.T410I (Heterozygous)	WT	p.E545K (Heterozygous)	WT	p.Q61H (Homozygous)	WT	WT
(NCI) H2228	WT	p.G31A (Heterozygous	WT	WT	WT	WT	WT	WT	p.H179R (Homozygous)
(NCI)H1568	WT	p.D77_E79del (Heterozygous)	WT	WT	WT	WT	WT	WT	p.Q331* (Homozygous)
(NCI)H1299	WT	WT	WT	WT	WT	WT	WT	WT	c.Del (Homozygous)
(NCI)H1395	WT	WT	WT	WT	WT	WT	WT	WT	WT

Table 5.2.3: Diverse genetic composition of the six commercial cell lines used to validate the bioinformatics analyses.

Somatic mutations in genes associated with lung cancer amongst the six commercial cell lines purchased from ATCC for validation of the bioinformatics analyses are indicated. The first column on the left-hand side shows the name of the cell line. Subsequent columns give details of the mutant gene (as shown at the top of the table), location of the mutation and whether the mutation is heterozygous or homozygous.

Cell line	Disease subtype
A549	ADC
(NCI)H460	Large cell carcinoma
(NCI)H2228	ADC
(NCI)H1568	ADC
(NCI)H1299	NSCLC lymphoid metastasis
(NCI)H1395	ADC

Table 5.2.4: Histological composition of the six commercial cell lines used to validate the bioinformatics analyses.

5.2.5 Lung cancer cell lines that harbour mutant *KEAP1* contain greater protein expression of NRF2-target genes than *NFE2L2* mutant cell lines.

To examine whether lung cell lines with mutant *KEAP1* overexpress NRF2-target genes to a greater extent than those with mutant *NFE2L2*, western blotting of proteins encoded by NRF2-target genes was carried out. Protein was extracted from the chosen representative *KEAP1*-mut (A549 and H460), *NFE2L2*-mut (H2228 and H1568) and “wildtype” (H1299 and H1395) cell lines. Aliquots (20 μg protein) of cell lysates were resolved according to molecular mass by SDS-PAGE, followed by western blotting for the protein expression of several NRF2-target genes. As shown in Figure 5.2.4, the two *KEAP1*-mut cell lines, show higher protein expression of varying degrees of the majority of the NRF2-target genes, when compared to the two *NFE2L2*-mut cell lines.

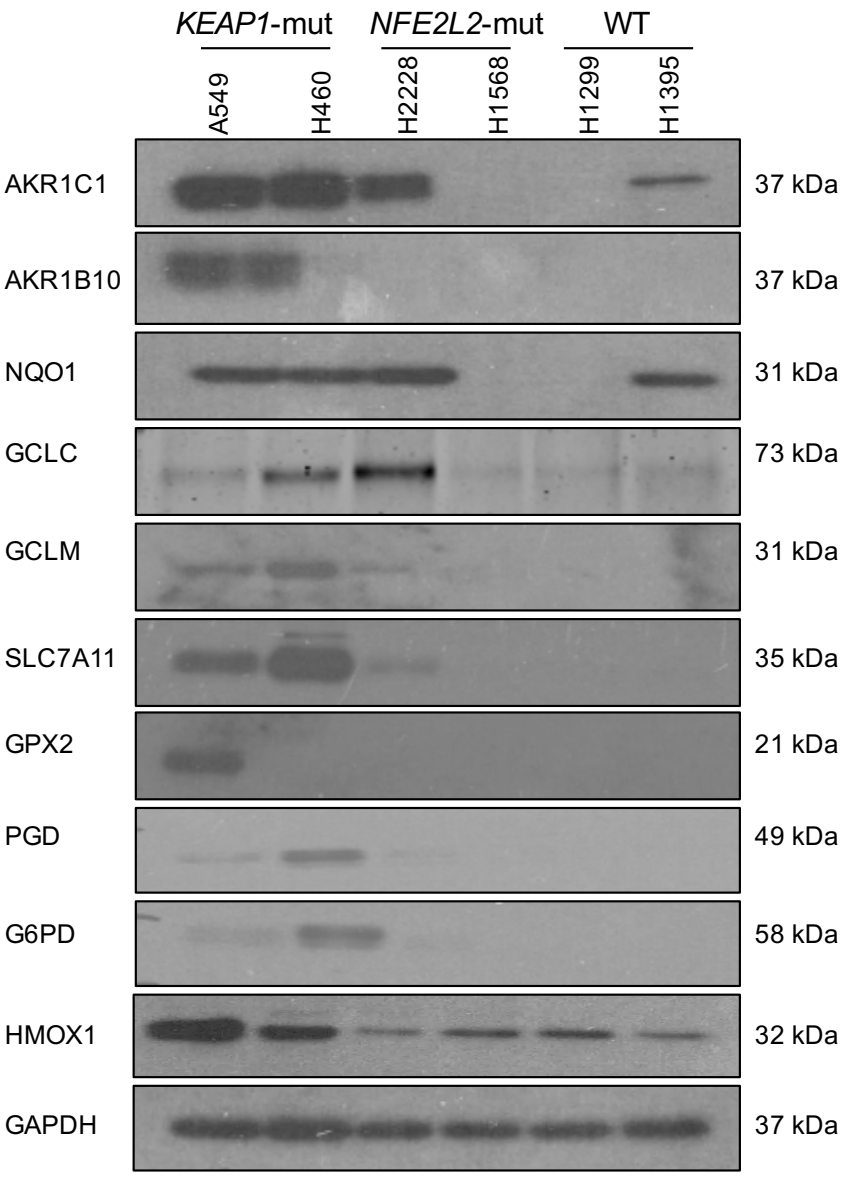


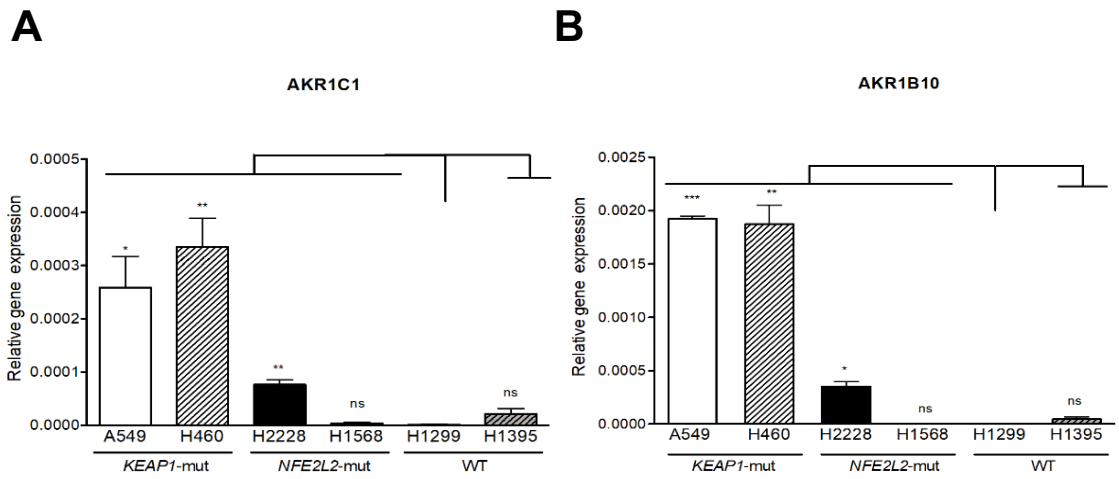
Figure 5.2.4: KEAP1 mutant lung cancer cell lines have greater protein expression of several NRF2-target genes in comparison to NFE2L2 mutant cell lines.

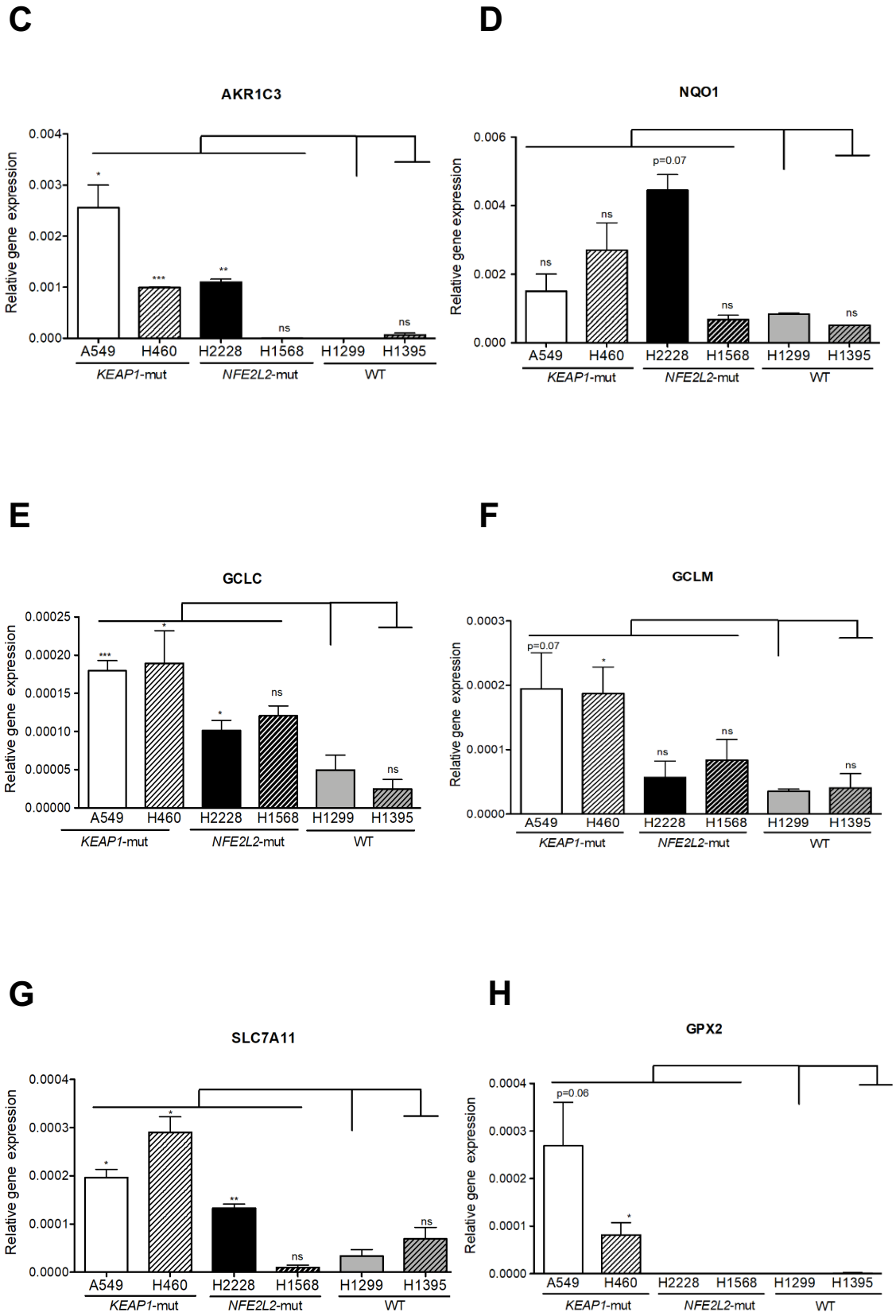
Protein expression analysis was carried out by western blotting. The name of each cell line is given at the top of the image, the name of the proteins being analysed is given down the left-hand side and their corresponding molecular weights down the right-hand side. The immunoblots data displayed are one representative image of three independent biological replicates.

5.2.6 *KEAP1*-mutant lung cancer cell lines express greater NRF2-target gene mRNA than do *NFE2L2*-mutant cell lines

To establish whether the increases in expression of proteins encoded by NRF2-target genes in *KEAP1*-mut and *NFE2L2*-mut cell lines, seen in Figure 5.2.5, are mirrored at a mRNA level. RNA was harvested from all six cell lines, then reverse transcribed into cDNA, before TaqMan qRT-PCR was carried out to access mRNA levels of several NRF2-target genes.

Figure 5.2.5 panels A-K reveal relatively higher levels of mRNA of NRF2-target genes in the two *KEAP1*-mut cell lines (A549 and H460) when compared with levels in the two *NFE2L2*-mut cell lines (H2228 and H1568). Surprisingly, NRF2 mRNA expression (shown in panel L) across the six cell lines does not correlate with the mRNA expression of the downstream NRF2-target genes. This, taken with the data shown in Figure 5.2.4, suggests that NRF2 mRNA itself it not a good marker of NRF2-target gene expression in lung cancer cell lines, presumably because it is principally controlled at the level of protein stability.




C
AKR1C3
Relative gene expression
0.004
0.003
0.002
0.001
0.000
*
***
**
ns
ns
A549
H460
H2228
H1568
H1299
H1395
KEAP1-mut
NFE2L2-mut
WT
D
NQO1
Relative gene expression
0.006
0.004
0.002
0.000
ns
ns
p=0.07
ns
ns
A549
H460
H2228
H1568
H1299
H1395
KEAP1-mut
NFE2L2-mut
WT
E
GCLC
Relative gene expression
0.00025
0.00020
0.00015
0.00010
0.00005
0.00000
***
*
*
ns
ns
A549
H460
H2228
H1568
H1299
H1395
KEAP1-mut
NFE2L2-mut
WT
F
GCLM
Relative gene expression
0.0003
0.0002
0.0001
0.0000
p=0.07
*
ns
ns
ns
A549
H460
H2228
H1568
H1299
H1395
KEAP1-mut
NFE2L2-mut
WT
G
SLC7A11
Relative gene expression
0.0004
0.0003
0.0002
0.0001
0.0000
*
*
**
ns
ns
A549
H460
H2228
H1568
H1299
H1395
KEAP1-mut
NFE2L2-mut
WT
H
GPX2
Relative gene expression
0.0004
0.0003
0.0002
0.0001
0.0000
p=0.06
*
A549
H460
H2228
H1568
H1299
H1395
KEAP1-mut
NFE2L2-mut
WT

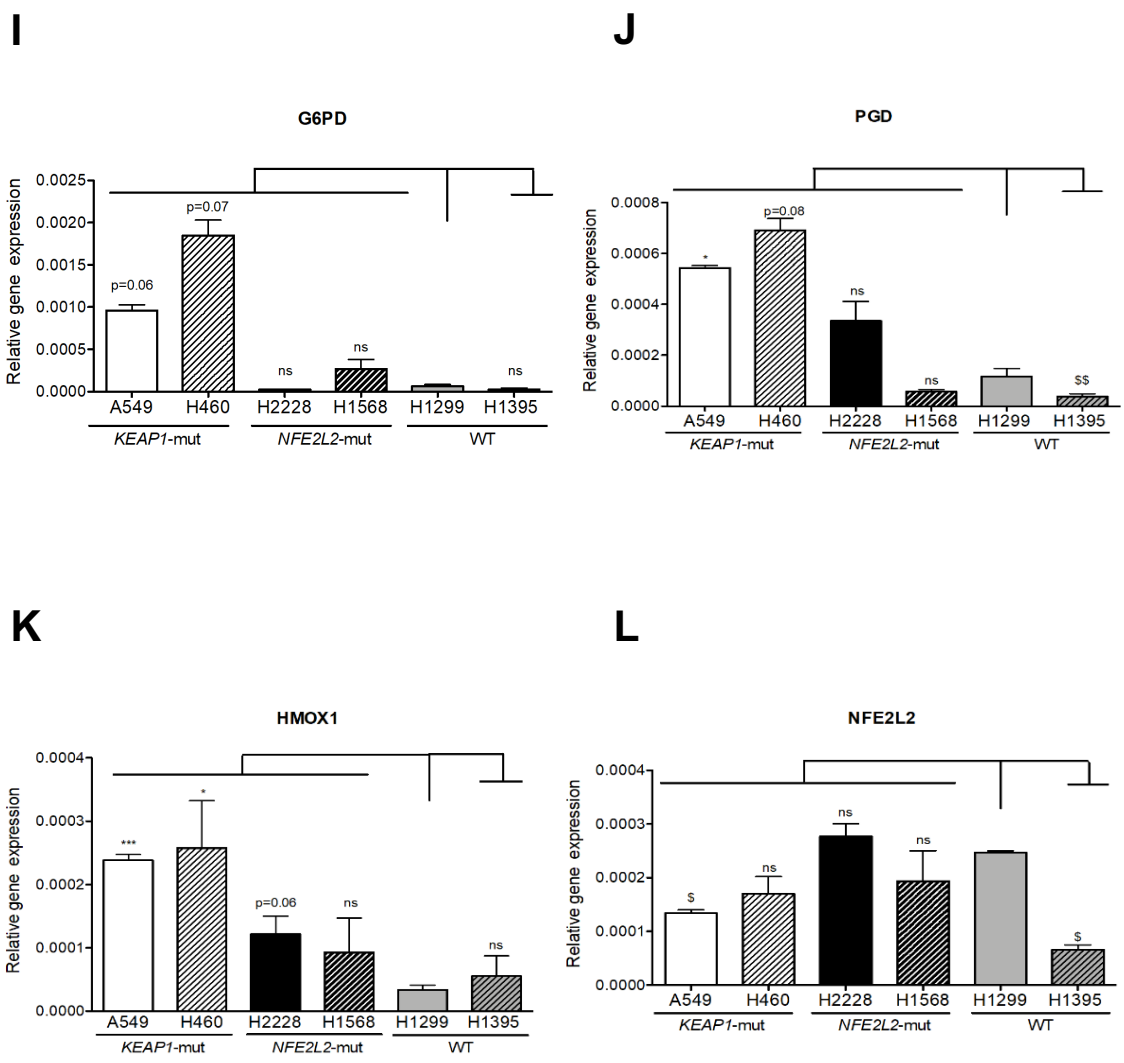


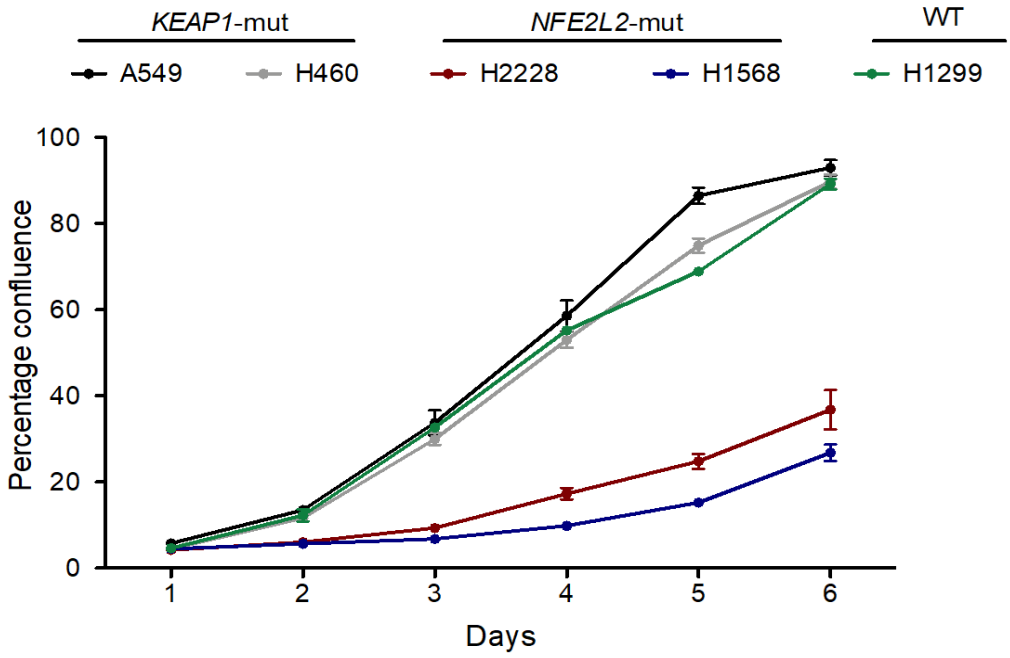
Figure 5.2.5: KEAP1 mutant cell lines express higher mRNA levels of NRF2-target genes than NFE2L2 mutant cell lines

*Analysis of mRNA expression across a panel of six cell lines. In every graph, the KEAP1-mut cell lines (A549 and H460) are represented in the solid white and adjacent black-striped bar, NFE2L2-mut cell lines (H2228 and H1568) are shown in the solid black and black-striped bars and KEAP1/NFE2L2 "wildtype" cell lines (H1299 and H1395) are shown in the grey and grey-striped bar. (A-K) mRNA expression of NRF2-target genes (L) NFE2L2 mRNA expression. All mRNA expression data presented was normalized to expression of the housekeeper 18S for each cell line. The mean of three independent experiments is plotted with the associated SEM value. T-tests were carried out to calculate significant changes in comparison to H1299 cells. All statistical analysis was carried out using GraphPad Prism 5 software. Significant increases are denoted with the * symbol; * represents P< 0.05, ** P<0.01 and ** P< 0.001. P values > 0.05 were deemed not significant and are denoted by "ns". Significant decreases are denoted with a $ symbol, and the same cut offs applied.*

5.2.7 *KEAP1* mutant cell lines show a greater proliferative ability than *NFE2L2* mutant cell lines.

The proliferative capacity of cell lines harbouring mutations in *KEAP1* or *NFE2L2* was accessed using the IncuCyte® Zoom imaging system and crystal violet clonogenicity assays. As shown in Figure 5.2.6A, A549 and H460 cell lines which both harbour mutations in *KEAP1* show a greater proliferation rate across six consecutive days than the H2228 and H1568 *NFE2L2*-mut cell lines. To access the colony formation ability of the six cell lines, each was seeded at the same low density and left to grow for two weeks, after which the colonies were fixed with ice-cold methanol before being stained with crystal violet. The A549 and H460 *KEAP1*-mut cell lines yield more stained colonies than either of the H2228 and H1568 *NFE2L2*-mut cell lines, Figure 5.2.6B. It is potentially not surprising that the two *KEAP1*-mut cell lines show enhanced proliferation as both of these cell lines also harbour additional mutations in *KRAS* which have been studied extensively for their ability to enhance cellular proliferation (Sunaga et al.,2011). Therefore, it is difficult at this time to attribute the highly proliferative capacity of the A549 and H460 cells solely due to their *KEAP1* mutation.

A



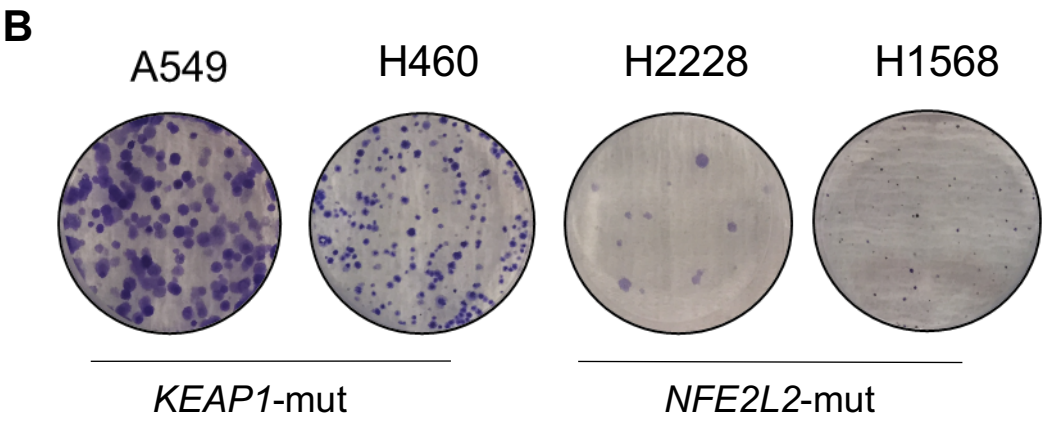


Figure 5.2.6: KEAP1-mut cell lines show higher proliferative capacity than NFE2L2-mut cell lines.

(A) Proliferation was assessed over six consecutive days using the IncuCyte® Zoom imaging system and percentage confluence was calculated using the associated software. Cells were seeded at a low density and plates were scanned every 24 hr for six consecutive days. In between scanning plates were maintained in a 37°C 5% CO_2 incubator. Data shown are one representative of three independent biological replicates, with the mean value (from 6 identically treated wells) and the associated SEM displayed on the graph. (B) To access colony formation a crystal violet colony stain was carried out. Cells were seeded at a low density and left to grow for two weeks. After which, plates were fixed with methanol and stained with crystal violet. The images shown are one representative of three independent biological replicates.

5.2.8 To eliminate the issue of the genetic diversity seen in the panel of six cell lines, a panel of isogenic CRISPR/Cas9 cell lines was generated

The six commercial cell lines used in the previous figures have a wide genetic diversity (as shown in Table 5.2.3) and this means it is not possible to attribute the changes in NRF2-signaling to just the presence of a *KEAP1* mutation or a *NFE2L2* mutation. Also, the data generated in the previous chapter highlighted that *KEAP1* and *NFE2L2* mutations tend to differ in their zygosity. Both A549 and H460 cells contain homozygous mutations in *KEAP1* and both H2228 and H1568 cells harbour heterozygous mutations in *NFE2L2*, meaning that we cannot directly compare these four cell lines. In an attempt to overcome this obstacle an isogenic panel of CRISPR/Cas9 cell lines was generated using the H1299 cell line background in order to mimic a *KEAP1* mutation or a *NFE2L2*

mutation, and so allow a direct comparison of their effects. The steps involved in the generation of the four cell CRISPR/Cas9 cell lines is shown in Figure 5.2.7. Guide RNA sequences were designed to target early in exon 1 of *KEAP1* and cloned into the LentiCRISPR-V2 backbone. The plasmid used to generate both the NRF2 knockout and NRF2 gain-of-function cell lines was kindly provided by Dr Laureano de la Vega (University of Dundee). Plasmids were then transfected into H1299 cells, 48 hr after which puromycin was added, to remove any non-transfected cells. After cells had recovered from puromycin selection to a confluency of approx. 80-90% in the well, the cells were trypsinized and seeded at a density of one cell per well in 96-well plates. The residual cell population was cultured in flasks and frozen down; this is the “pooled” population. The 96-well plates were then routinely screened by eye to ensure the growth of only one colony per well. Once the colony had reached approx. 80-90% confluency in the well, it was trypsinised and expanded to a 24-well plate. After the cells in the 24-well plate grew to a confluency of 80-90%, they were trypsinised and 90% of the total cells were taken for mRNA and protein analyses. The remaining 10% of cells were transferred to a 12-well plate and left to grow until the protein/mRNA data was obtained for that clone. Positive clones, showing the desired alterations in both mRNA and protein expression, were expanded and frozen down for long term storage at -80^{o}C. The details of the four cell lines generated is given in Table 5.2.5. An empty vector cell line (EV H1299) was made through the transfection of an empty LentiCRISPR-V2 plasmid containing no guide RNA sequence. A *KEAP1* knock out cell line, referred to as KKO H1299, was generated to mimic a cell line processing a mutation in *KEAP1*. A *NFE2L2* gain-of-function cell line, referred to as GOF H1299 was generated to mimic a cell line with a mutation in the Neh2 domain of NRF2. Also, a *NFE2L2* knock out cell line, referred to as NKO H1299, was generated to ensure that all of the parameters of NRF2-signaling that will be accessed across the panel are dependent on NRF2.

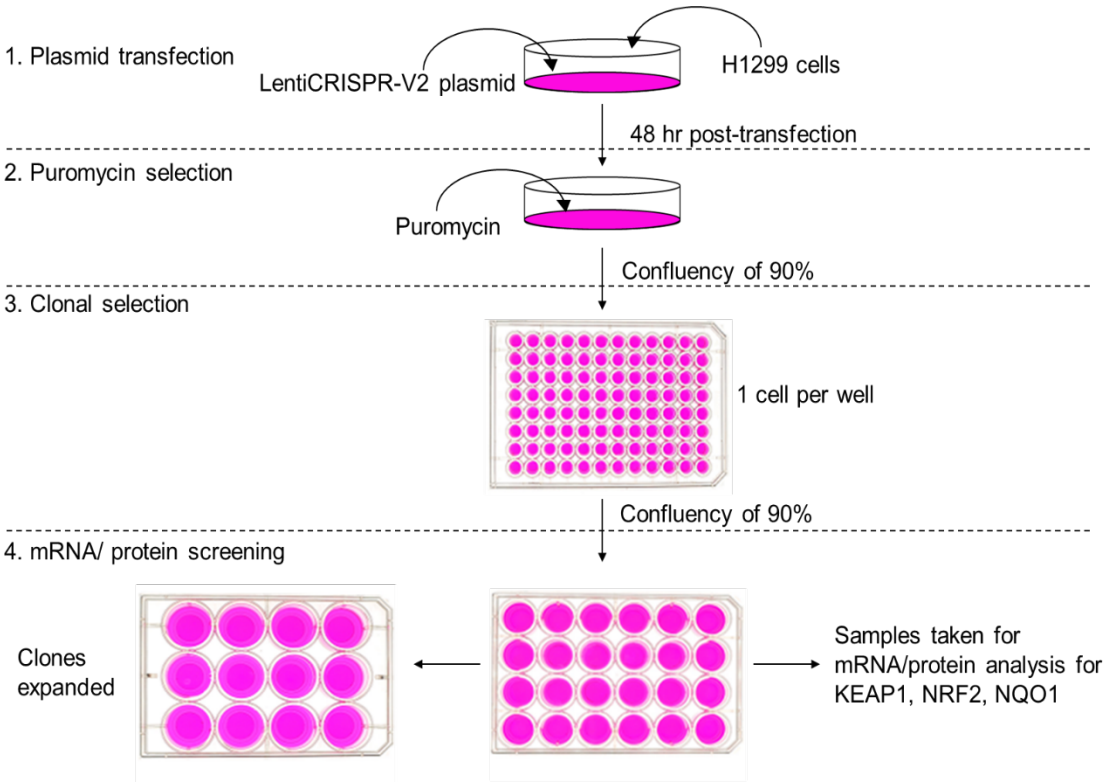


Figure 5.2.7: Diagram of the step-wise generation of the CRISPR/Cas9 cell lines

The four stages taken to generate the panel of four CRISPR/Cas9 cell lines are highlighted down the left-hand side of the image

Cell line name	Details of CRISPR alteration	Level of NRF2
EV H1299	Control cell line transfected with empty Lenti-CRISPR V2	-
KKO H1299	Knockout of *KEAP1* (mimicking a *KEAP1* mutant cell line)	High
GOF H1299	Gain-of-function due to disruption of the Neh2 domain of NRF2 (mimicking a *NFE2L2* mutant cell line)	High
NKO H1299	Knockout of *NFE2L2* (functions as a control)	Low

Table 5.2.5: Details for the four CRISPR/Cas9 generated cell lines

Table explaining the cell line name, details of the four CRISPR/Cas9 cell lines that were generated using the parental H1299 cell line and the corresponding predicted level of NRF2.

5.2.9 Identification of H1299-derived CRISPR/Cas9-manipulated cell lines with altered NRF2 activity.

To be sure that the CRISPR/Cas9 protocol was successful and that we had generated the desired cell lines, expression of several NRF2-target genes was assessed at both an mRNA and a protein level, across several clones.

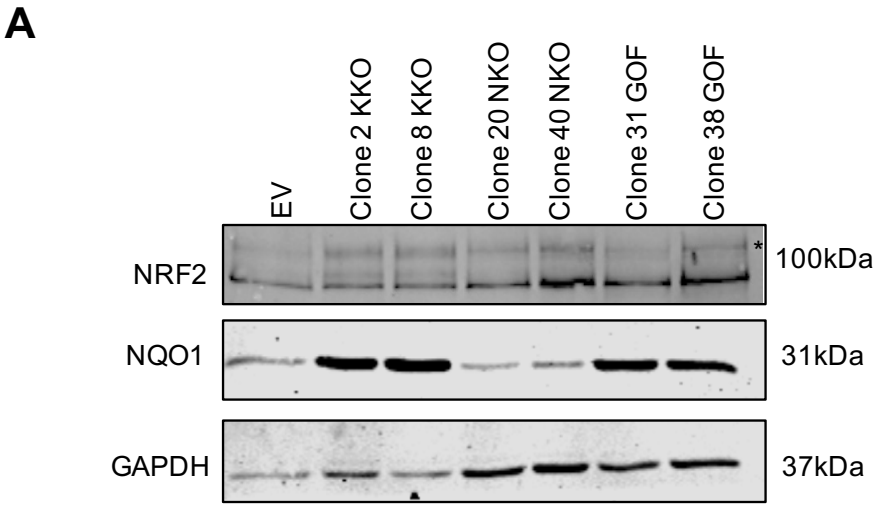
Figure 5.2.8A show the changes in protein expression and the densitometry, across just two clones per cell line. On average, between 30 and 40 clones were analysed per cell line. As assessed by western blotting, clone 2 KKO and clone 8 KKO both contained *KEAP1* knockout cells due to their increased abundance of NRF2 and NQO1 proteins when compared with the empty vector H1299 cell line (EV). Also, as assessed by western blotting, clone 20 and clone 40 NKO had reduced levels of NQO1 protein when compared with the EV cell line indicating they contain NRF2 knockout cells; however, NRF2 protein levels were not clearly altered in these cells when compared to the control. Both clone 31 GOF and clone 38 GOF were found to have higher NQO1 protein levels when compared with the EV control cell lines suggesting the cells have enhanced NRF2 activity.

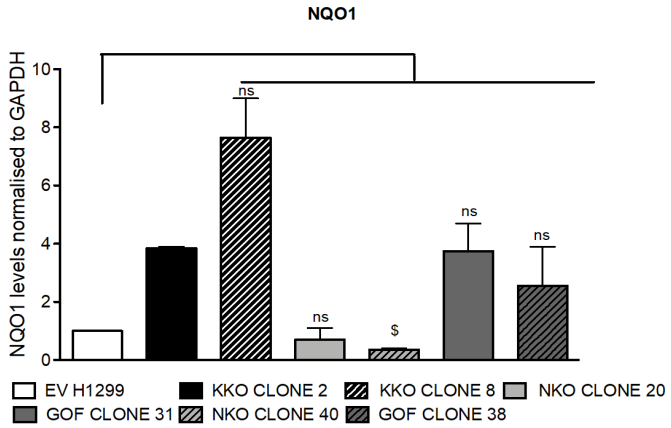
TaqMan mRNA analysis was carried out to determine whether the expression of *NFE2L2, KEAP1* and *NQO1* was appropriately changed in the cell lines, as shown in Figure 5.2.8B. When compared to the EV H1299 cell line, both clone 2 and clone 8 KKO have enhanced expression of *NQO1*, consistent with them representing *KEAP1* knockout cells. Clone 20 and clone 40 NKO cell lines show reduced expression of *NQO1* in comparison to the EV control cell line, consistent with them representing NRF2 knock out cells. By contrast, clone 31 and clone 38 GOF have elevated *NQO1* mRNA levels when compared to the EV H1299 cells, consistent with them containing NRF2 gain-of-function cells.

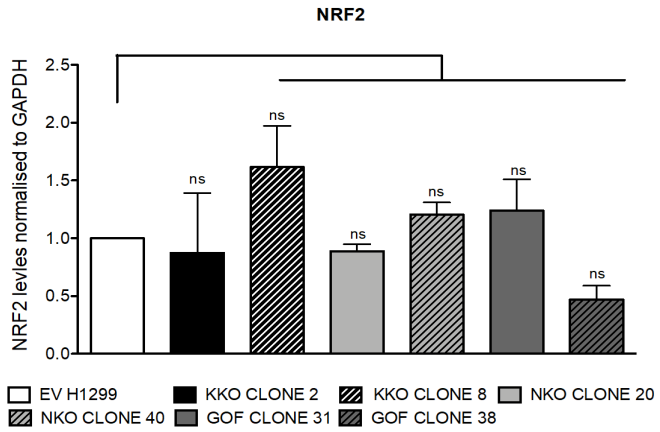
From the above analyses, clone 8 KKO, clone 40 NKO and clone 31 GOF were selected for further study, and the rest of the experiments described in this chapter used these cell lines, referring to them as KKO, NKO and GOF, respectively.

At the time that these analyses were carried out we did not have an antibody that could detect KEAP1. However, we did repeat the KEAP1 protein

analysis across the three chosen clonal cell lines at a later date; the data that are subsequently obtained is shown in Appendix 8.8.







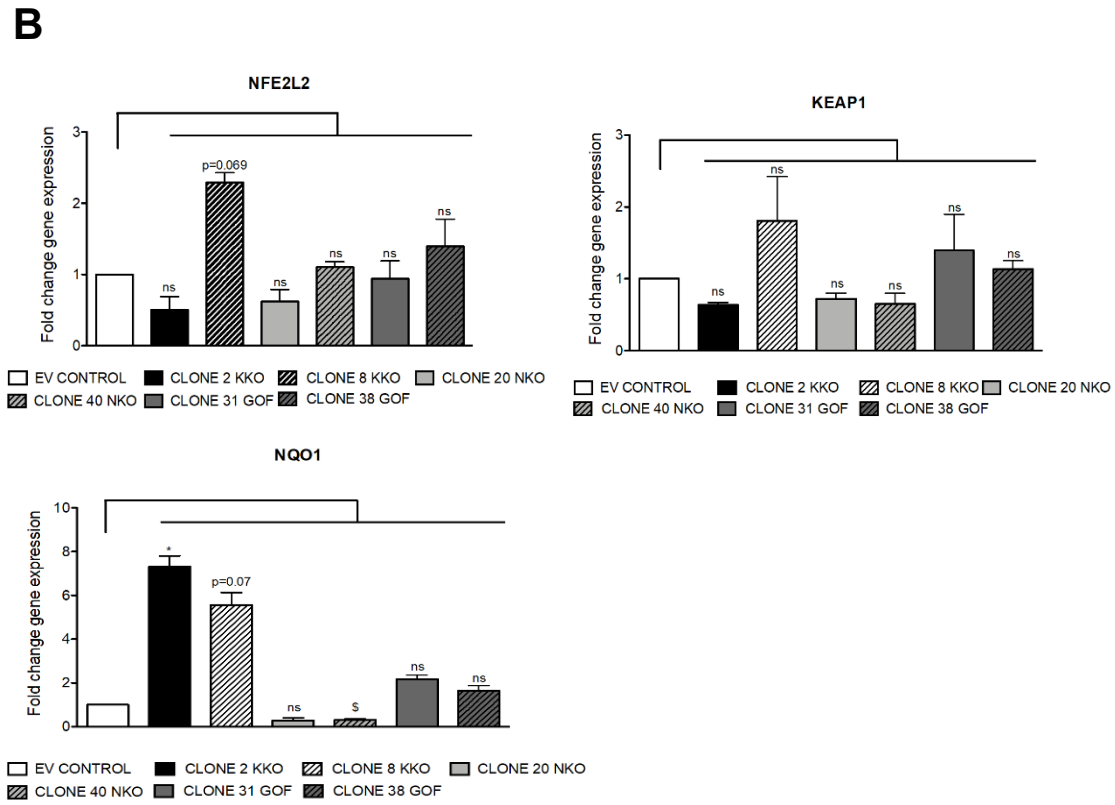


Figure 5.2.8: Confirmation via protein and mRNA analyses of the generation of four H1299-derived CRISPR/Cas9 cell lines

*(A) The levels of NRF2 (the band corresponding to NRF2 is indicated with a * symbol), NQO1 and GAPDH proteins were analysed across the EV, two positive KEAP1-knockout clones (clone 2 KKO and clone 8 KKO), two positive NRF2-knockout clones (clone 20 NKO and clone 40 NKO) and two positive NRF2 gain-of-function (GOF) clones (clone 31 GOF and clone 38 GOF). This image is representative of three independent biological replicates. (B) Densitometry of (A) normalized to GAPDH (house keeper) protein expression was carried out using the In-built LI-COR odyssey software. The mean protein expression value from three independent biological replicates is plotted with the associated SEM. Statistically significant changes relative to the EV cell line are displayed on each graph. A T-test was carried out using GraphPad Prism software to calculate significant alterations in expression, significant increases are marked with a * and significant decreases with a dollar sign. (C) The mRNA levels of NFE2L2, KEAP1 and NQO1 were analysed across the same cell lines as in (A). The data displayed are gene expression normalized to the expression of the house keeper protein, actin. The mean value from three independent biological replicates in plotted with the associated SEM. All significant increases are denoted with the * symbol; * represents P< 0.05, ** P<0.01 and ** P< 0.001. P values > 0.05 were deemed not significant and are denoted by "ns". Significant decreases are denoted with a $ symbol, and the same cut offs applied.*

5.2.10 KEAP1-knockout H1299 cells show higher levels of nuclear NRF2 than do NRF2-gain-of-function H1299 cells

NRF2 translocates into the nucleus to bind ARE sequences in the regulatory regions of its target genes and so increase their transcription. To determine whether KKO cells have higher levels of active nuclear NRF2, in comparison to GOF cells, western blotting of subcellular fractions was carried out.

All four EV, KKO, NKO and GOF H1299 CRISPR/Cas9 cell lines were seeded in large dishes. Once the cells reached a confluency of approx. 80-90% they were collected and fractionated through a series of centrifugation steps with two buffers. Aliquots (30 μg of protein) from both the cytoplasmic (C) and nuclear (N) fractions for each cell line were then loaded onto a SDS-PAGE gel and western blotting was carried out. As shown in Figure 5.2.9A, KKO cells have higher levels of nuclear NRF2 in comparison to both the EV and GOF cell lines. This may in part account for the considerably higher NQO1 expression in the KKO cells when compared to the GOF cells as more nuclear NRF2 suggests more active transcription of its target genes. To determine whether the nuclear accumulation of NRF2 in the KKO cell line could be further enhanced through inhibition of $SCF^{\beta\text{-}TrCP}$-mediated degradation of NRF2, cells were treated with the GSK-3 inhibitor CT99021 for 24 hr prior to fractionation. The nuclear levels of NRF2 do not appear to be altered by CT99021 treatment in any of the four CRISPR cell lines (Figure 5.2.9B). To confirm that this was not due to incomplete inhibition of GSK-3, before fractionation 20% of the treated whole-cell lysate was retained and immunoblotted for GSK-3 and pGSK-3 Ser9/21. In both Figure 5.2.10A and B Lamin A/C has been included as a positive control for nuclear fractionation and alpha-tubulin is used as a positive control for the cytoplasmic fraction.

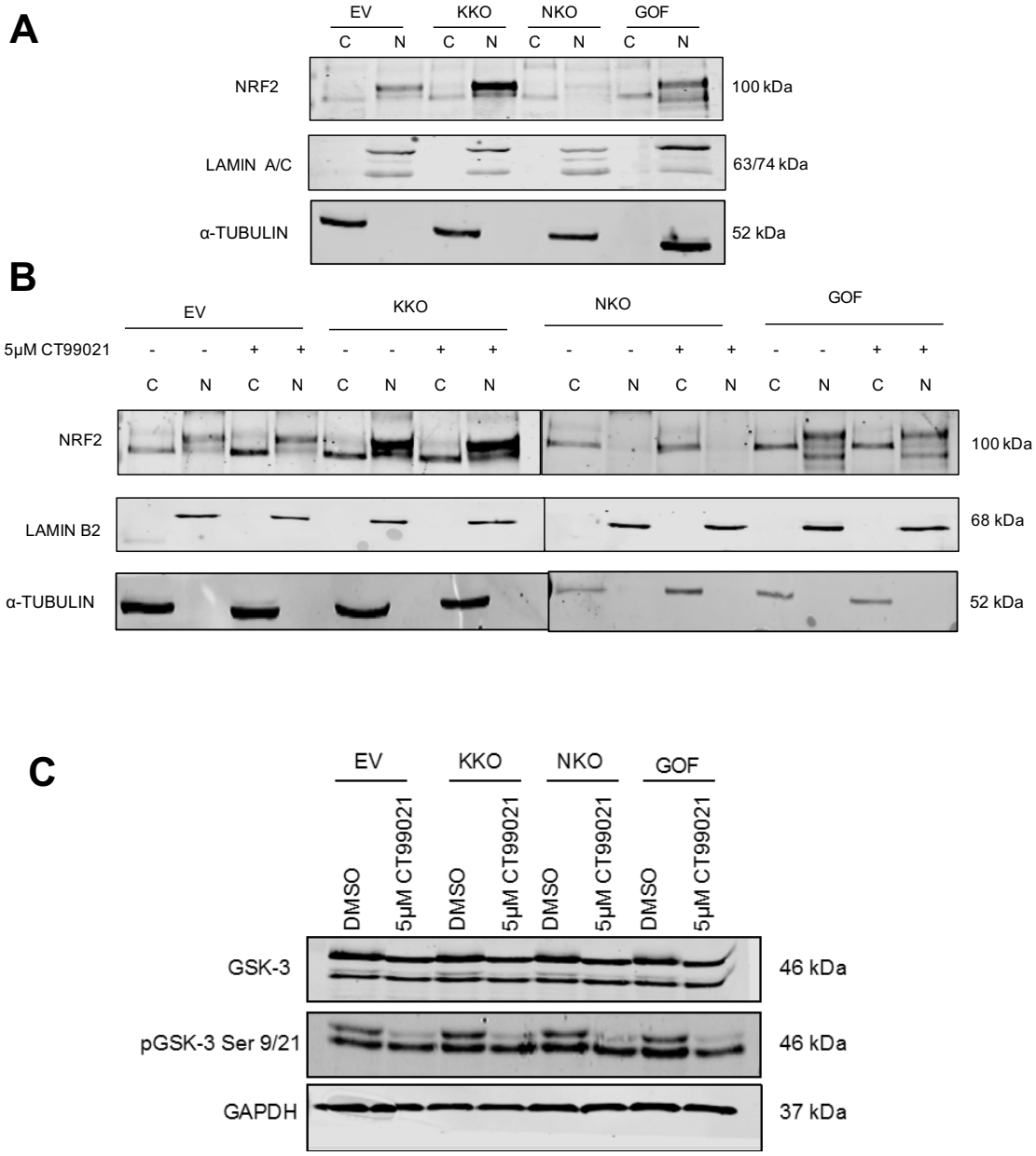
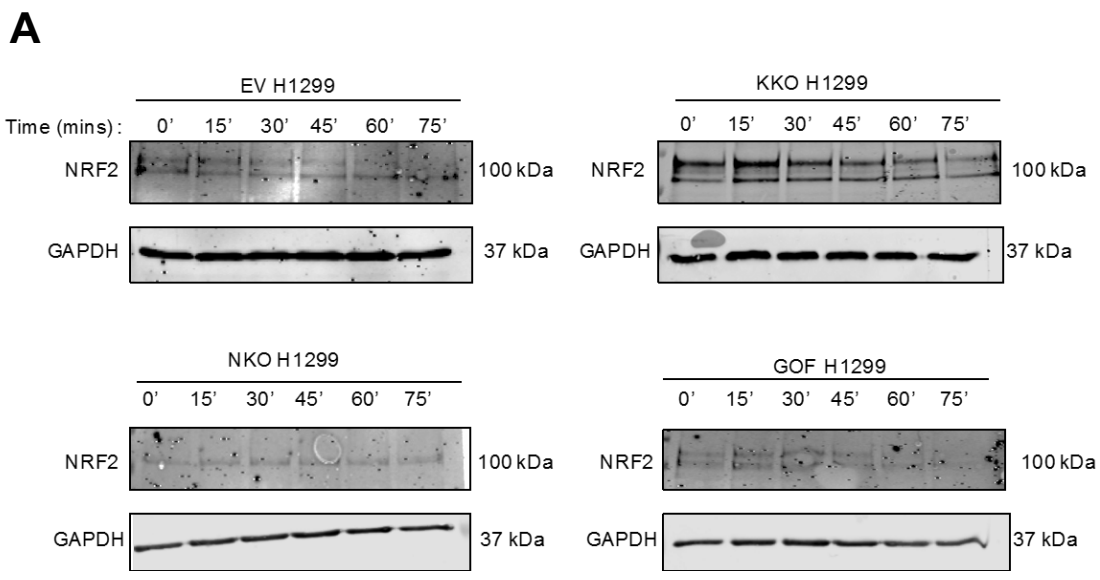


Figure 5.2.9: KEAP1-knockout H1299 cells show higher nuclear NRF2 protein levels than NRF2-gain-of-function H1299 cells.

(A) Subcellular fractionation of the four CRISPR/Cas9 cell lines empty vector (EV), KEAP1 knock out (KKO), NRF2 knock out (NKO) and NRF2 gain of function (GOF). This image is representative of images obtained from three independent biological replicates. (B) Cells were treated with either DMSO or 5 μM CT99021 for 24 hr before fractionation. This image one representative image of three independent biological replicates. (C) To confirm that treatment with CT99021 leads to an inhibition of GSK-3, all four CRISPR/Cas9 cell lines were treated with either DMSO of 5 μM CT99021 for 24 hr before checking protein expression via western blot. This image is from one experiment and is representative of three independent biological replicates.

5.2.11 The half-life of endogenous NRF2 protein is greater in KKO H1299 cells than in GOF H1299 cells.

Endogenous NRF2 protein is relatively unstable with a half-life of approximately 30 min in resting cells (McMahon et al., 2003). A longer NRF2 half-life would suggest that degradation of the transcription factor was less efficient and would indicate higher NRF2 activity. To determine whether the half-life of NRF2 protein differs in the KKO and GOF cell lines, a CHX chase experiment was carried out; CHX is a potent inhibitor of protein synthesis, which can be used to assess protein stability. Cells were treated with CHX for either: 0, 15, 30, 45, 60 or 75 min before cells were collected and protein extracted. Portions (30 μg of protein) of cell lysate were loaded on SDS-PAGE and western blotting was carried out to detect NRF2 protein levels. As shown in Figure 5.2.10A and B, the half-life of endogenous NRF2 protein in KKO H1299 cells is considerably longer than in GOF H1299 cells, with calculated half-lives of 63 min and 35 min respectively.



B

Cell line				
	EV H1299	KKO H1299	NKO H1299	GOF H1299
Half-life of endogenous NRF2 (min)	24.8	63.5	-	34.7
Fold change relative to EV H1299 cells	-	2.6	-	1.4

Figure 5.2.10: The half-life of endogenous NRF2 protein is greater in KEAP1-knockout H1299 cells than in NRF2-gain-of-function H1299 cells

(A) CHX chase was used to determine the half-life of endogenous NRF2 protein in the four CRISPR/Cas9 cell lines. This image is representative of three independent biological replicates. (B) Table showing the calculated half-life of endogenous NRF2 across the four CRISPR cell lines. The half-life of endogenous NRF2 in the NKO H1299 cells could not be calculated due to the low protein expression.

5.2.12 The activity of NQO1 is higher in KKO H1299 cells than in GOF H1299 cells.

To determine if NQO1 activity is higher in the KKO cells than in the GOF cells, an adapted version of the Prochaska NQO1 enzyme assay was carried out (Fahey et al., 2004). The data displayed in Figure 5.2.11, reveal that KKO cells (shown in the black bar) have enhanced NQO1 activity in comparison to the GOF cells (shown in the dark grey bars) at every concentration of CDDO-ME. Not surprisingly, treatment with CDDO-ME does not appear to increase NQO1 activity in the KKO cells. Possibly more surprisingly, CDDO-ME also failed to increase NQO1 activity in GOF cells. This is likely to be because CDDO-ME functions to inhibit KEAP1-mediated ubiquitination of NRF2. Reassuringly, it was found that NQO1 activity was increased by CDDO-ME in EV H1299 cells in a dose-dependent manner, and as expected the NQO1 activity in the NKO cell line was found to be unaffected by CDDO-ME treatment.

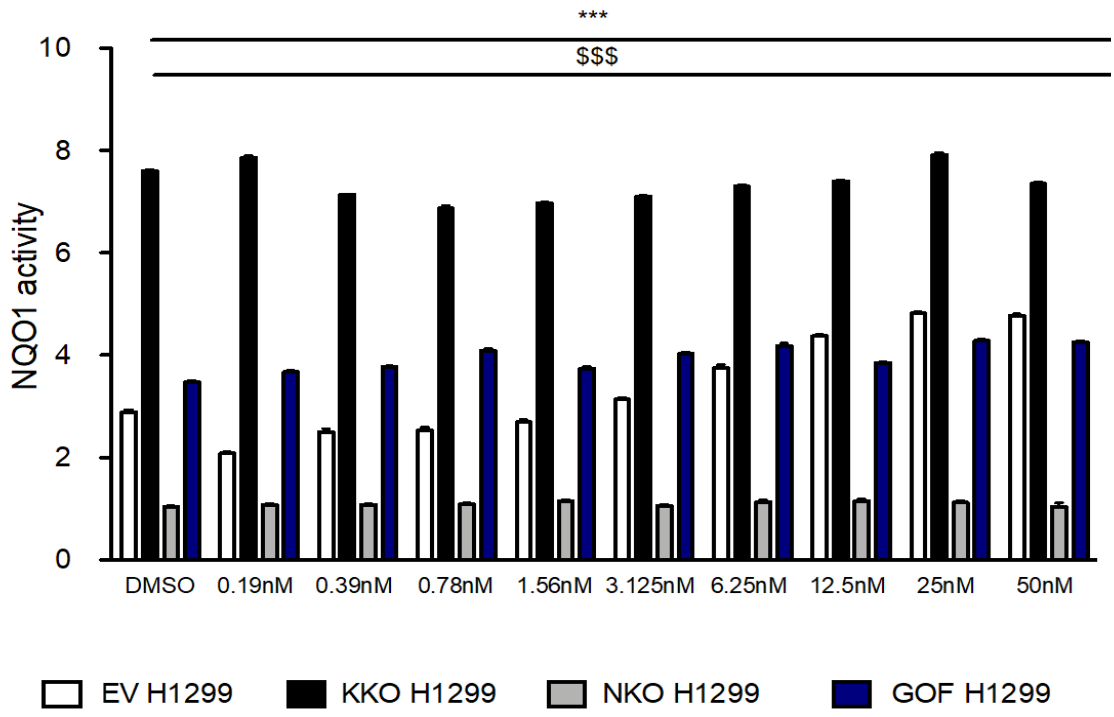


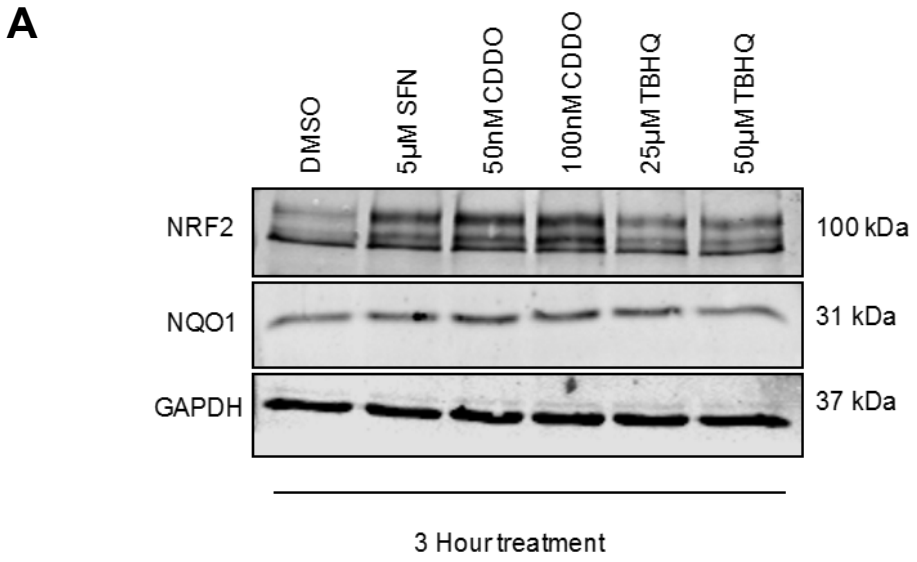
Figure 5.2.11: KEAP1-knockout H1299 cells have significantly greater NQO1 activity in than NRF2-gain-of-function H1299 cells

*The enzymatic activity of NQO1 across the four CRISPR/Cas9 cell lines was determined following treatment with increasing doses of an inducing agent. Bars on the graph represent the mean value normalized to total protein concentration and associated SEM. T-tests were carried out to calculate significant changes in comparison to H1299 cells at each concentration of CDDO-ME. All statistical analysis was carried out using GraphPad Prism 5 software. All significant increases are denoted with the * symbol; * represents P< 0.05, ** P<0.01 and ** P< 0.001. P values > 0.05 were deemed not significant and are denoted by "ns". Significant decreases are denoted with a $ symbol, and the same cut offs applied. The *** indicates that the KKO cells had statistically higher NQO1 activity in comparison to the EV cells in the presence of every concentration of CDDO-ME. The $$$ indicates that the NKO cells have statistically lower NQO1 activity in comparison to the EV at each concentration of NRF2-activator.*

5.2.13 Optimization for the use of an NRF2 activator.

To gain a greater understanding about the alterations in NRF2-target gene expression across the CRISPR panel, an NRF2 activator was selected. Firstly, as a control, “wild-type” EV H1299 cells were treated for either 3 hr (shown in Figure 5.2.12A) or 24 hr (shown in Figure 5.2.12B) with either SFN, CDDO-ME, or tBHQ; DMSO was used as a vehicle control. At the 3-hr time point, an increase in NRF2 protein levels, relative to the DMSO vehicle control, was observed following treatment with SFN, both concentrations of CDDO-ME and both concentrations of tBHQ. The greatest increase in NRF2 protein levels at this time point was seen with 100 nM CDDO-ME. The 3-hr time point was too early to see changes in levels of NQO1 protein level.

At the 24-hr time point all inducers at all concentrations increased NRF2 protein levels relative to the control. All NRF2-activators increased NQO1 protein expression after treatment for 24 hr, with the most pronounced increase seen with 50 nM CDDO-ME. Considering the data from both Figure 5.2.12A and B, 50n M CDDO-ME was chosen to be used in the following experiments.



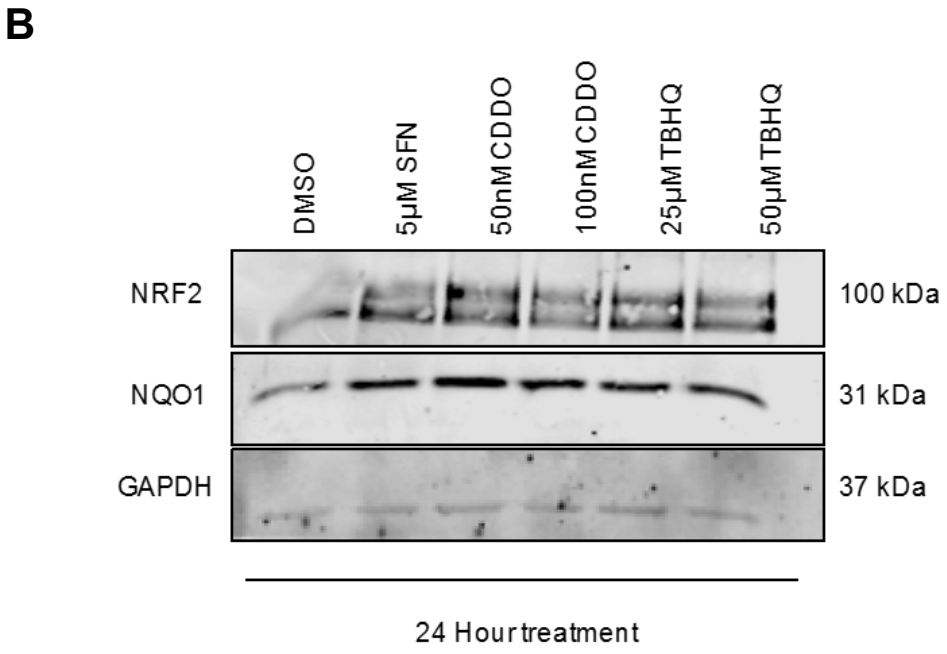


Figure 5.2.12: Treatment of "wild-type" EV H1299 cells with inducers increases the abundance of NRF2 protein and NQO1

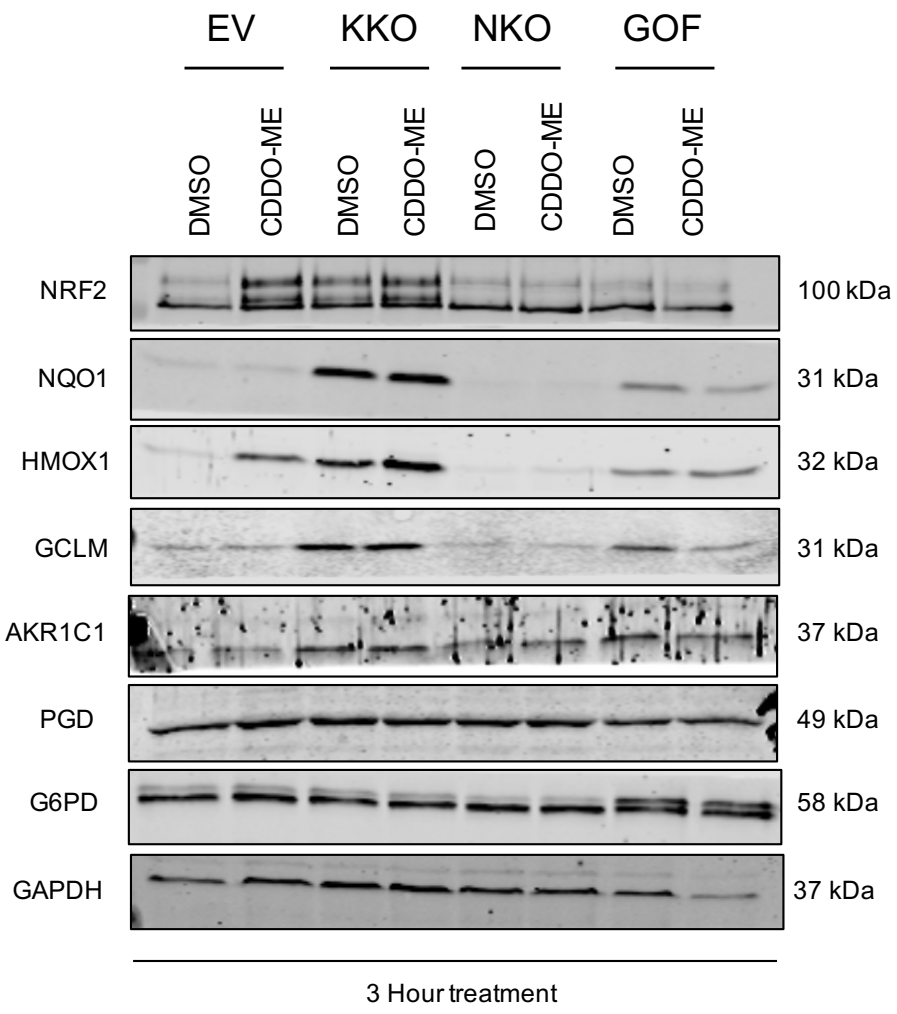
*(A) EV H1299 cells, which contain wild-type KEAP1 and NRF2, were treated with inducers for 3 hr before NRF2 and NQO1 protein levels were measured by western blot. The band on the image corresponding to NRF2 protein is indicated with a * symbol. This image represents three independent biological replicates. (B) EV H1299 cells, were treated with inducers for 24 hr before NRF2 and NQO1 protein levels were measured. All treatments were carried out in DMEM supplemented with 10% FCS. This image is one representative image of three independent biological replicates.*

5.2.14 KKO H1299 cells have comparability greater protein expression of NRF2-target genes than GOF H1299 cells.

To validate data from the bioinformatics analyses (Figure 5.2.2), the levels of protein encoded by several NRF2-target genes in both KKO and GOF H1299 cell lines that had been treated with DMSO l or CDDO-ME for 3 or 24 hr were assessed. As shown in Figure 5.2.13A, KKO cells have higher basal protein levels of NRF2, NQO1, HMOX1, and GCLM, than the GOF H1299 cells. At the 24-hr time point, shown in Figure 5.2.13B, CDDO-ME treatment increases the abundance of NRF2, NQO1, HMOX1 and GCLM proteins in the EV H1299 cell line relative to the control. By contrast, in the KKO H1299 cell line, CDDO-ME

treatment surprisingly only increased the protein levels of HMOX1 relative to the control. Potentially NRF2, NQO1 and GCLM proteins may be at their highest detectable levels in the KKO H1299 cells already and therefore further increases are not possible. Furthermore, NQO1 protein levels appear to be unaffected by CDDO-ME treatment in GOF cells, whereas HMOX1 protein is increased by CDDO-ME in these cells. Even after treatment with CDDO-ME, NQO1, HMOX1 and GCLM protein levels were not higher in the GOF cells than in untreated KKO cells. These findings suggest that removal of *KEAP1* has a more pronounced effect on NRF2-target gene expression that disruption of the KEAP1-binding ETGE motif in NRF2.

A



B

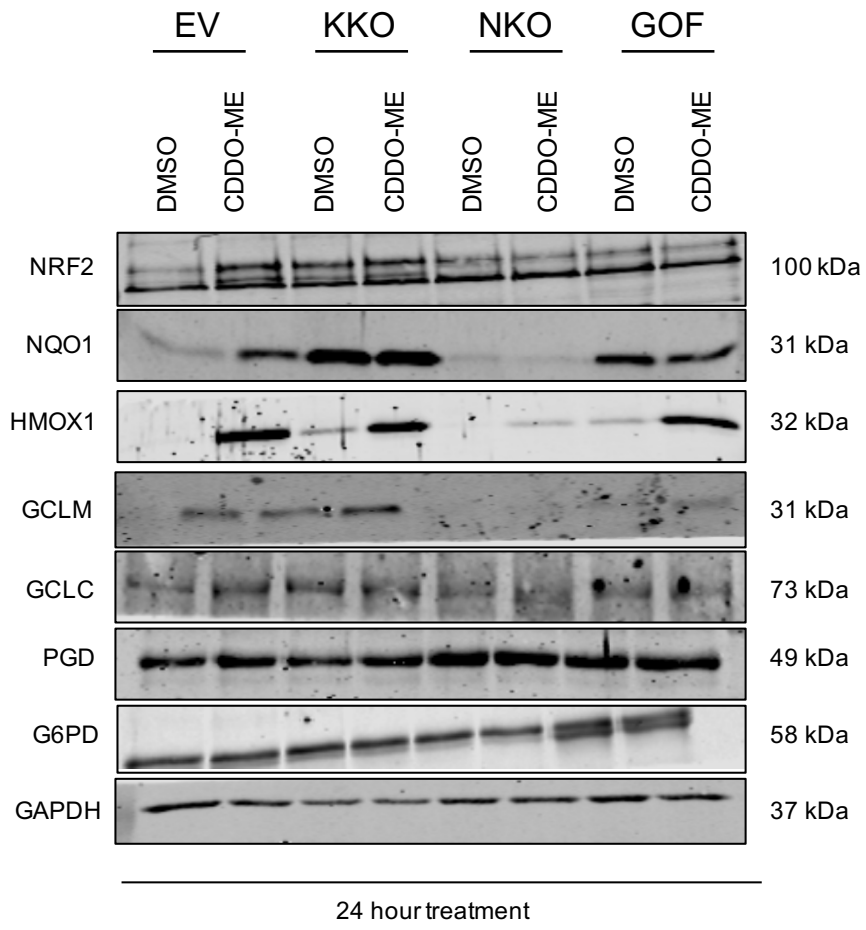


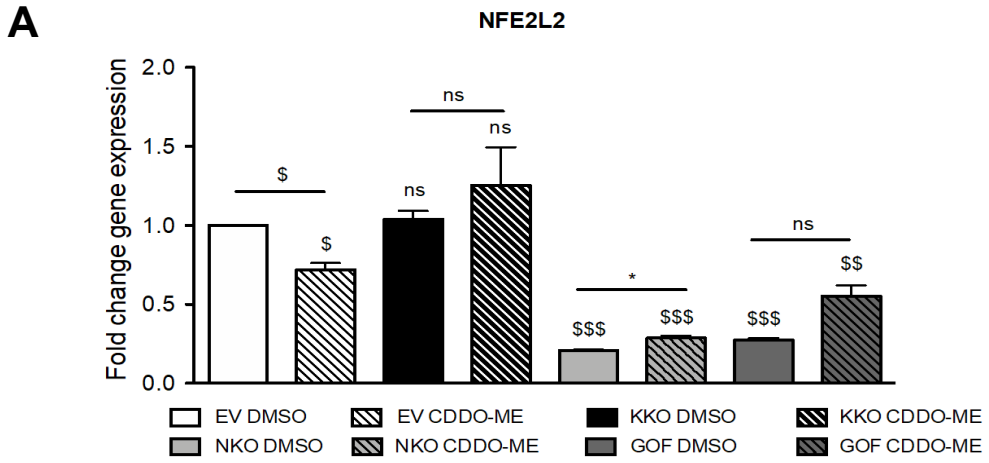
Figure 5.2.13: KKO H1299 cells contain higher levels of proteins encoded by NRF2-target genes than do GOF H1299 cells.

(A) Protein analysis of several NRF2-target genes across the panel of four CRISPR/Cas9 cell lines: EV, KKO, NKO and GOF, treated either with DMSO or 50 nM CDDO-ME for 3 hr. Cell lines and treatments are indicated across the top of the image. The names of the proteins being analysed are given down the left-hand side and their corresponding molecular weights down the right-hand side of the image. This image is one representative image of three independent biological replicates. (B) Protein analysis of several NRF2-target genes across the panel of four CRISPR/Cas9 cell lines: EV, KKO, NKO and GOF treated either with DMSO or 50 nM CDDO-ME for 24 hr. Cell lines and treatments are indicated across the top of the image. The names of the proteins being analysed are given down the left-hand side and their corresponding molecular weights down the right-hand side of the image. All treatments were carried out in DMEM supplemented with 10% FCS. This image is one representative image of three independent biological replicates.

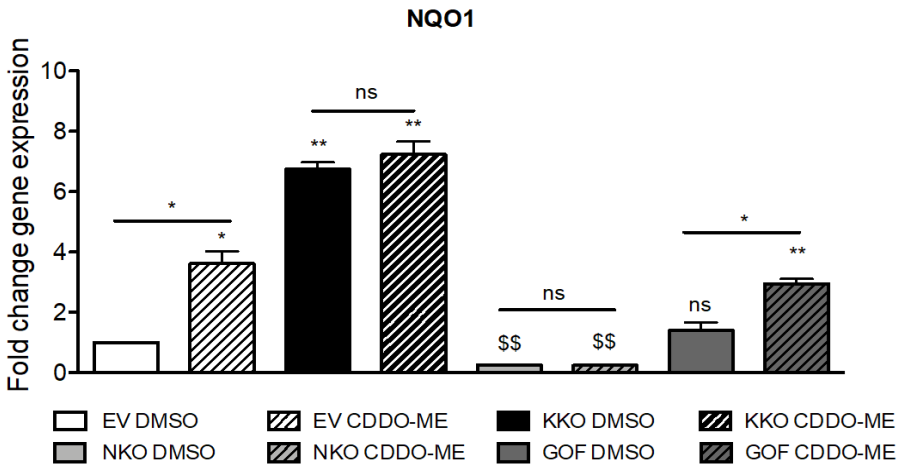
5.2.15 Expression of mRNA for NRF2-target genes is higher in KKO H1299 cells than GOF H1299 cells

To establish whether the increase in abundance of proteins encoded by NRF2-target genes in KKO H1299 cells (Figure 5.2.13) is reflected at a mRNA level, analysis of several NRF2-target genes was carried out across the CRISPR/Cas9 panel of cells that had been treated with either 50 nM CDDO-ME or DMSO for 24 hr.

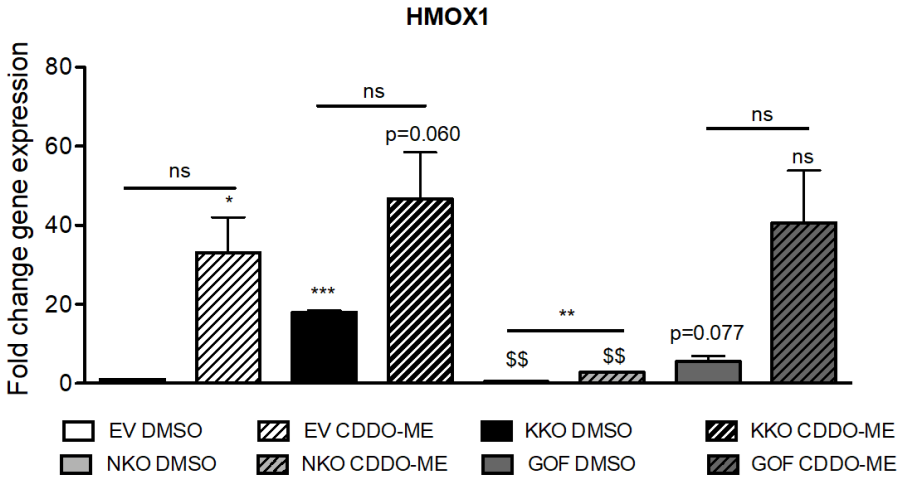
As shown in Figure 5.2.14 A-F, the levels of mRNA for *NQO1, SLC7A11, GCLC, GLCM* and *HMOX1* were increased by treatment of the EV, KKO and GOF cell lines with 50 nM CDDO-ME for 24 hr, whereas treatment with CDDO-ME did not significantly alter the expression of *NFE2L2* across the panel of CRISPR/Cas9 cell lines. Even in the presence of CDDO-ME the GOF H1299 cells still showed lower mRNA expression of several NRF2-target genes when compared to untreated KKO H1299 cells. This further supports the idea that knockout of KEAP1 has a great effect on NRF2-signalling that a disruption of KEAP1-mediated degradation.



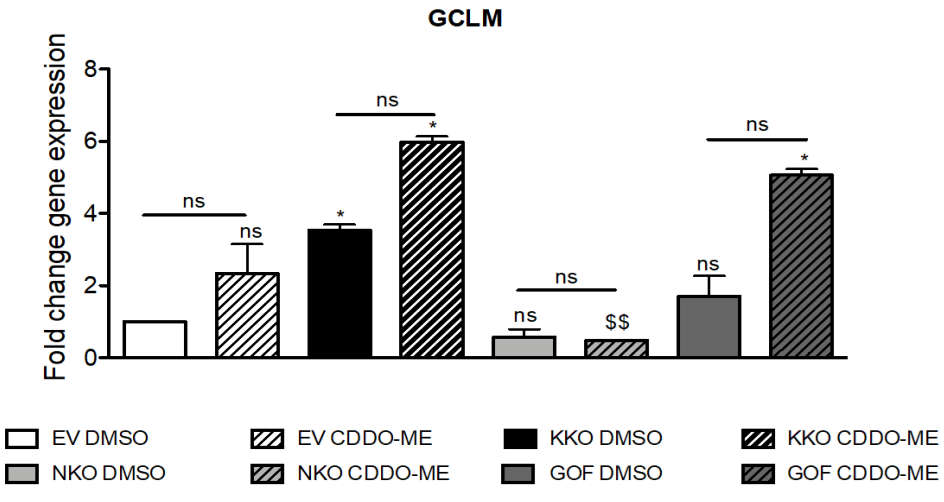
B


NQO1
Fold change gene expression
10
8
6
4
2
0
ns
**
**
*
*
ns
$$
$$
*
**
ns
EV DMSO
EV CDDO-ME
KKO DMSO
KKO CDDO-ME
NKO DMSO
NKO CDDO-ME
GOF DMSO
GOF CDDO-ME

C


HMOX1
Fold change gene expression
80
60
40
20
0
ns
*
ns
***
p=0.060
**
$$
$$
ns
p=0.077
ns
EV DMSO
EV CDDO-ME
KKO DMSO
KKO CDDO-ME
NKO DMSO
NKO CDDO-ME
GOF DMSO
GOF CDDO-ME

D


GCLM
Fold change gene expression
8
6
4
2
0
ns
ns
*
ns
*
ns
ns
$$
ns
ns
*
EV DMSO
EV CDDO-ME
KKO DMSO
KKO CDDO-ME
NKO DMSO
NKO CDDO-ME
GOF DMSO
GOF CDDO-ME

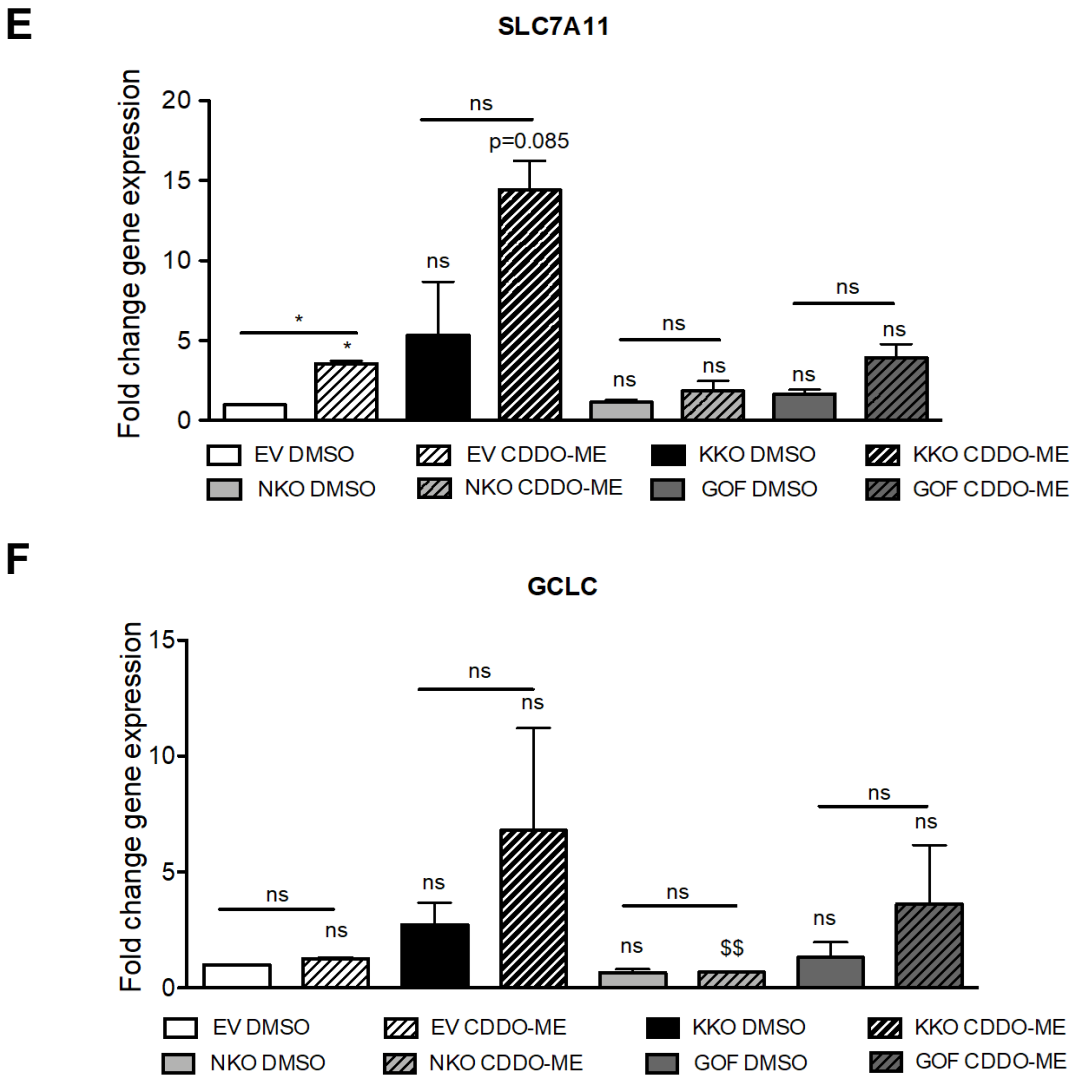


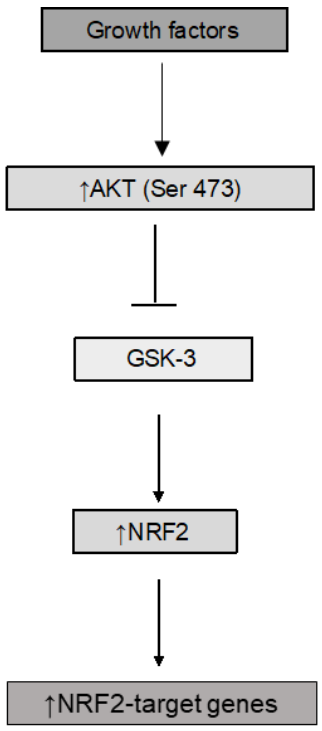
Figure 5.2.14: KKO H1299 cells constitutively overexpress NRF2-target genes to a greater extent than EV H1299 or GOF H1299 cells, and the NRF2-target genes in KKO H1299 cells can be induced by CDDO-ME.

*(A-F) The expression of NRF2-target genes was measured across the panel of H1299 CRISPR cell lines using TaqMan qRT-PCR. Cells were either treated with DMSO or 50 nM CDDO-ME for 24 hr in DMEM supplemented with 10% FCS. The data displayed are fold change gene expression normalized to the expression of the house keeper protein, actin. The value on each graph represents the mean value obtained from three independent biological replicates and the associated SEM. T-tests were calculated using GraphPad Prism software. Significant alterations relative to the EV H1299 DMSO-treated control are indicated on the graphs. Also, significant alterations between the DMSO treated and CDDO-ME treated bars for that cell line is indicated above the black line. All significant increases are denoted with the * symbol; * represents P< 0.05, ** P<0.01 and ** P< 0.001. P values > 0.05 were deemed not significant and are denoted by "ns". Significant decreases are denoted with a $ symbol, and the same cut offs applied.*

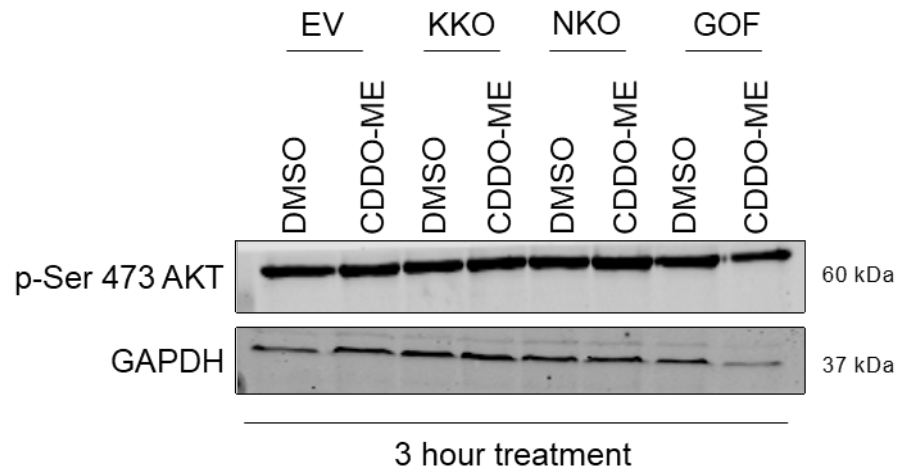
5.2.16 The enhanced expression of NRF2-target genes observed in KEAP1-knockout cells relative to NRF2-gain-of-function cells is not due to alterations in AKT signalling.

Active AKT signalling has been shown to phosphorylate GSK-3 at Ser-9 and Ser-21 leading to its inactivation (Cross et al., 1995). When GSK-3 is inactivated it can no longer phosphorylate NRF2, thus halting $SCF^{\beta\text{-}TrCP}$-mediated degradation and increasing NRF2-target gene expression. Thus, changes in AKT signalling could potentially account for the greater NRF2-target gene expression seen in the KKO cells compared to that in GOF cells (Figure 5.2.15A). To determine whether KKO cells have higher AKT signalling than the GOF cells, western blotting was carried out. As shown in Figure 5.2.15B and C, neither treatment with CDDO-ME for 3 or 24 hr altered the protein expression of p-Ser 473 AKT across any of the four CRISPR/Cas9 cell lines. Also, the basal expression of p-Ser 473 AKT was not altered between the KKO and GOF cells. This suggests that the alterations in NRF2-target gene expression seen between the KKO and the GOF cell line is not due to differences in AKT signalling. However, to be certain, the protein expression of p-Thr 308 and total AKT should be determined, and the phosphorylation at both sites should be calculated relative to the total AKT level.

A



B



C

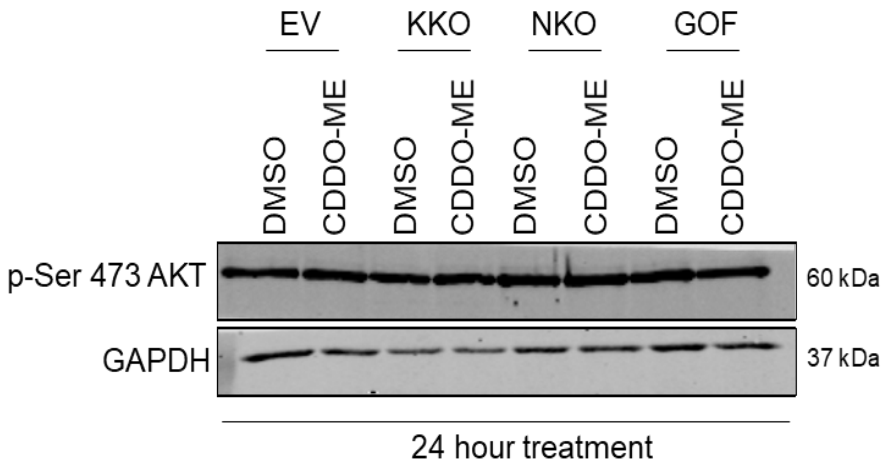


Figure 5.2.15: Changes in NRF2-target gene expression between the H1299 CRISPR cell lines is not due to differences in AKT signalling

(A) Diagram showing how the phosphorylation of AKT affects NRF2 activity and NRF2-target gene expression. (B) Protein analysis of AKT activity across the H1299 panel treated with either DMSO or 50 nM CDDO-ME for 3 hr in 10% FCS supplemented DMEM. The cell line name and treatments are indicated above the image. Protein expression of p-Ser 473 was used as a read out of AKT activity. This image is one representative image of three independent biological replicates. (C) Protein analysis of AKT activity across the H1299 panel treated with either DMSO or 50 nM CDDO-ME for 24 hr in10% FCS supplemented DMEM. The cell line name and treatments are indicated above the image. Protein expression of p-Ser 473 was used as a read out of AKT activity. This image is one representative image of three independent biological replicates.

5.2.17 Inhibition of AKT does not affect the levels of protein encoded by NRF2-target genes in any of the four CRISPR cell lines

To determine whether inhibition of AKT would lead to a reduction in NRF2-target gene expression, western blotting was carried out using lysates treated with the AKT inhibitor, MK2206. The data shown in Figure 5.2.16 demonstrates that neither treatment with 1 μM MK2206 or 5 μM MK2206 for 24 hr alters: NRF2, NQO1 or HMOX1 protein levels. This could be due to MK2206 not fully inactivating AKT, so the abundance of p-Ser 473 AKT was examined. Reassuringly the protein expression of p-Ser 473 AKT is decreased in every cell line treated with MK2206 at both concentrations.

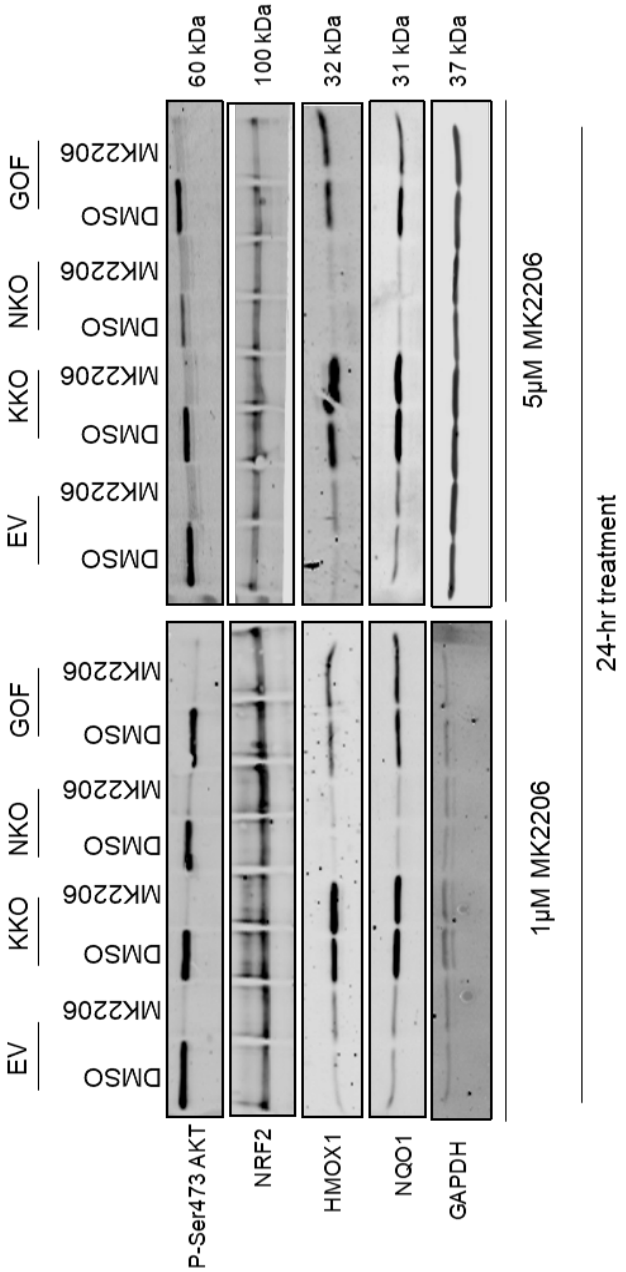


Figure 5.2.16: Treatment with the AKT inhibitor MK2206 does not affect NRF2-target gene expression across the four CRISPR/Cas9 cell lines

EV, KKO, NKO and GOF H1299 cells were treated with either 1 μM or 5 μM MK2206 for 24 hr in 10% FCS supplemented DMEM. Western blotting was then carried out to assess the protein expression of NRF2, NRF2-target genes and p-Ser 473 AKT. This image is one representative image of three independent biological replicates.

5.2.18 The expression of other KEAP1 binding proteins is not altered between the KKO and GOF cell lines.

Work by Ben Major and colleagues has identified several KEAP1-interacting proteins that can function as indirect regulators of NRF2 through their abilities to compete for KEAP1 binding (Hast et al., 2013). To determine if the expression of KEAP1 interacting proteins is different in the GOF and KKO cells, the protein levels of: TSC22D4, NRF1, p62/SQSTM1 and PGAM5 were measured by western blotting. As shown in Figure 5.2.17, the protein levels of none of these four KEAP1-interacting proteins were altered in any of the four CRISPR/Cas9 cell lines. It should be noted the protein expression of p62 is slightly reduced in the NKO H1299 cell line, this is to be expected as p62 is a NRF2-target gene.

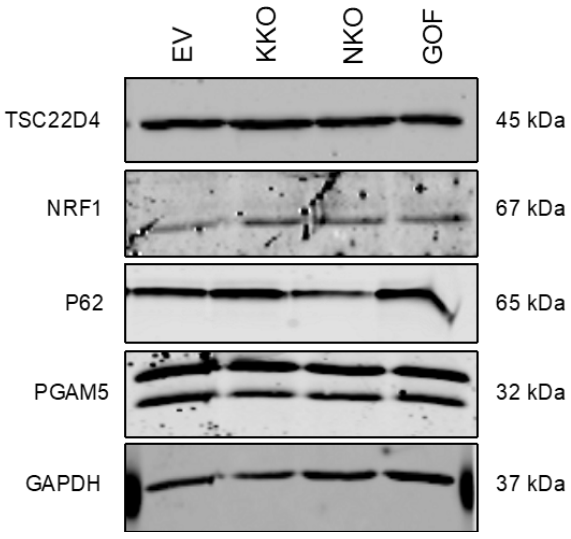


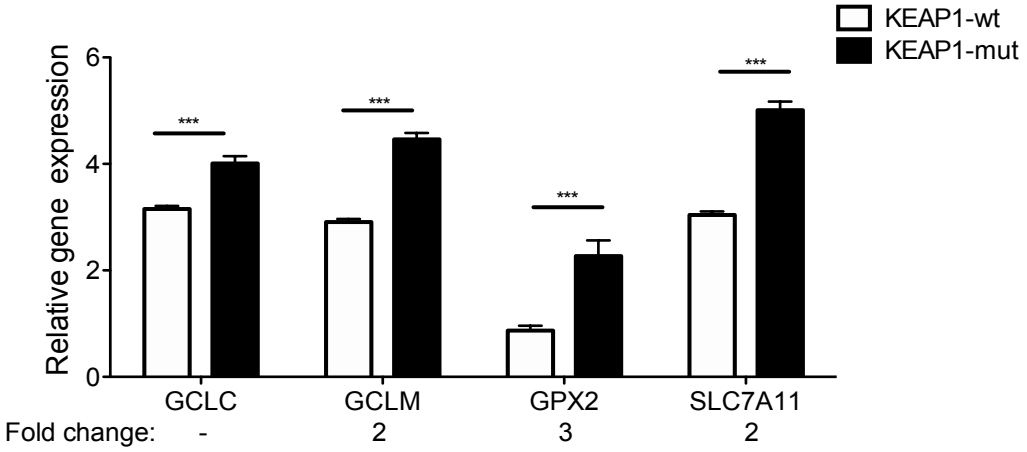
Figure 5.2.17: The abundance of the KEAP1-binding proteins TSC22D4, NRF1, PGAM and p62 is not substantially altered by knockout of KEAP1.

The protein expression levels of four KEAP1-interacting proteins that may out compete NRF2 for binding to KEAP1 was compared across the four CRISPR/Cas9 H1299 cell lines by western blotting. This image is one representative image of three independent biological replicates.

5.2.19 Glutathione-associated NRF2-target genes are overexpressed to a greater extent in *KEAP1* mutant cells than in *NFE2L2* mutant cells.

In the bioinformatics analyses described in Section 5.2.2 the expression of NRF2-target genes associated with the synthesis and metabolism of glutathione, in both *KEAP*1-mut and *NFE2L2*-mut lung cancer cell lines, was investigated. Using the Sanger, AstraZeneca, and CCLE datasets and the software TIBCO Spotfire®, lung cancer cell lines (239 in total) were filtered for the presence of somatic mutations in either *KEAP1* or *NFE2L2*. Next, the expression of *GCLC, GCLM, SLC7A11* and *GPX2* between the *KEAP1*-mut (28 cell lines) and *KEAP1*-wt (211 cell lines) cell populations (as shown in Figure 5.2.18A) and the *NFE2L2*-mut (8 cell lines) and *NFE2L2*-wt (231 cell lines) cell line populations (Figure 5.2.18B) was examined. This revealed the *KEAP1*-mut cell lines have a significantly higher expression of all four GSH-associated NRF2-target genes when compared with *KEAP1*-wt cell lines. Also, the *NFE2L2*-mut cell lines have significantly higher expression of *GPX2* when compared to the *NFE2L2*-wt cell lines. The expression of *GCLC, GCLM* and *SLC7A11* is not significantly altered between cell lines harbouring a mutation in *NFE2L2* or wildtype for the gene. Overall these data suggest that *KEAP1*-mut cell lines express NRF2-target genes associated with glutathione production and expenditure, to a greater extent than cell lines harbouring mutations in *NFE2L2*.

A



B

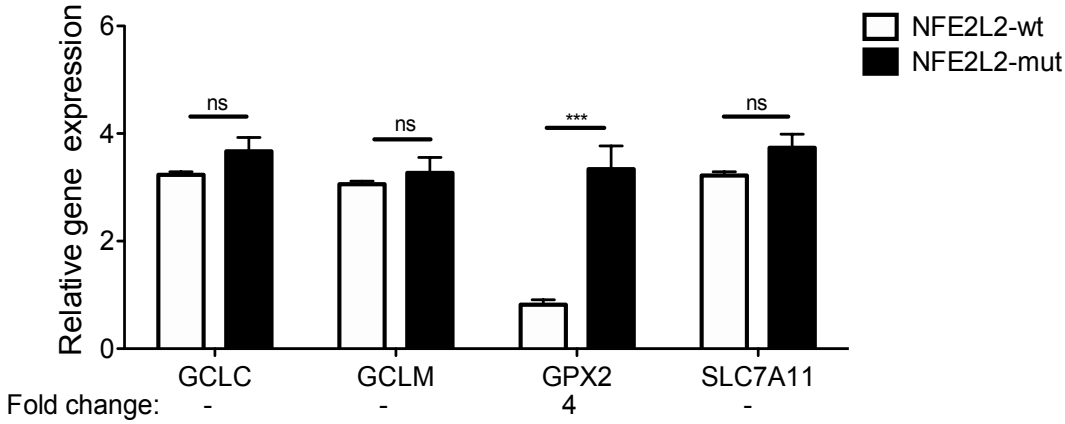


Figure 5.2.18: GSH-associated NRF2-target genes are more highly expressed in KEAP1 mutant cell lines than in NFE2L2 mutant cell lines.

*The black bars on each graph represent the mutant cell line population and the white bars represent the wildtype population. The names of all four genes analysed is given on the x-axis. The fold change in gene expression of the mutant cell lines relative to the wildtype counterparts is displayed underneath the graph. Statistically significant change in gene expression in the mut cell lines relative to the wt cell lines are indicated on each graph. (A) Analysis of the expression of four NRF2-target genes associated with GSH in lung cancer cell lines that are either KEAP1-mut (28 cell lines) or KEAP1-wt (211 cell lines). (B) Analysis of the expression of four Nrf2-target genes associated with GSH in lung cancer cell lines that are either NFE2L2-mut (8 cell lines) or NFE2L2-wt (231 cell lines). Data presented on both graphs represents the mean value obtained, and the error plotted is the associated SEM. All statistical analysis was carried out using GraphPad Prism 5 software. All significant increases are denoted as follows: *, P< 0.05; **, P<0.01; *** P< 0.001. P values > 0.05 were deemed not significant and are denoted by "ns". Significant decreases are denoted with a $ sign.*

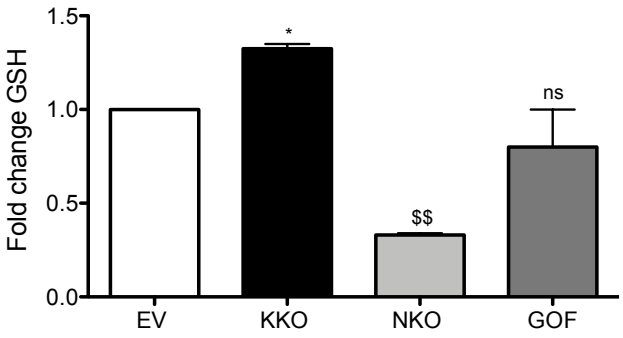
5.2.20 KKO H1299 cells contain higher levels of glutathione than GOF H1299 cells.

The relative overexpression of *GCLC, GCLM* and *SLC7A11* in *KEAP1*-mut cells suggests they ought to contain higher glutathione levels. To investigate this possibility, total glutathione was measured across the four CRISPR H1299 cell lines using the mCB probe. mCB is a cell-permeable nonfluorescent compound that fluoresces upon contact with thiols such as glutathione. As shown in Figure 5.2.219A, KKO H1299 cells contain significantly higher concentrations of total glutathione than either EV H1299 or GOF H1299 cells. NKO H1299 cells contain significantly lower glutathione levels when compared to the control cell line. Surprisingly GOF H1299 cells do not show a significant alteration in glutathione levels in comparison to the EV H1299 cells. However, it should be noted that there is a large associated error bar with the GOF H1299 data which may account for this.

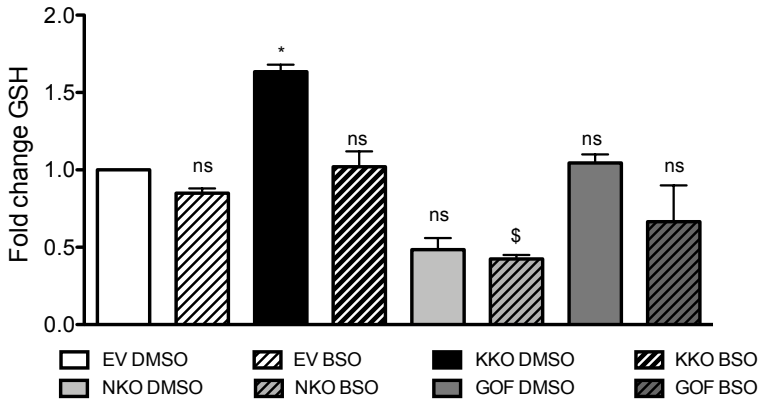
BSO is a potent inhibitor of glutamate-cysteine ligase that decreases intracellular glutathione concentrations. As shown in Figure 5.2.19B, treatment with BSO reduces the total glutathione level in KKO H1299 cells and GOF H1299 cells relative to the DMSO control for each cell line. With the changes being significant in the KKO H1299 cells but only tending towards significance in the GOF H1299 cells.

NAC is an antioxidant that stimulates glutathione production in the cell. Treatment with NAC could only increase total glutathione levels in the EV H1299 cell line and not in either the KKO, NKO or GOF cell lines (Figure 5.2.19C). In the case of the KKO H1299 cells this may be due to them already having a very high level of glutathione that cannot be further increased. NKO H1299 cells have no NRF2, and since NRF2 controls the expression of genes involved in the production of glutathione, it makes sense that this pathway cannot be induced in these cells through NAC treatment. However, it is unclear at this moment to why the levels of glutathione in the GOF H1299 cells are not affected by NAC administration.

A


Fold change GSH
1.5
1.0
0.5
0.0
*
ns
$$
EV
KKO
NKO
GOF

B


Fold change GSH
2.0
1.5
1.0
0.5
0.0
*
ns
ns
ns
$
ns
ns
EV DMSO
EV BSO
KKO DMSO
KKO BSO
NKO DMSO
NKO BSO
GOF DMSO
GOF BSO

C

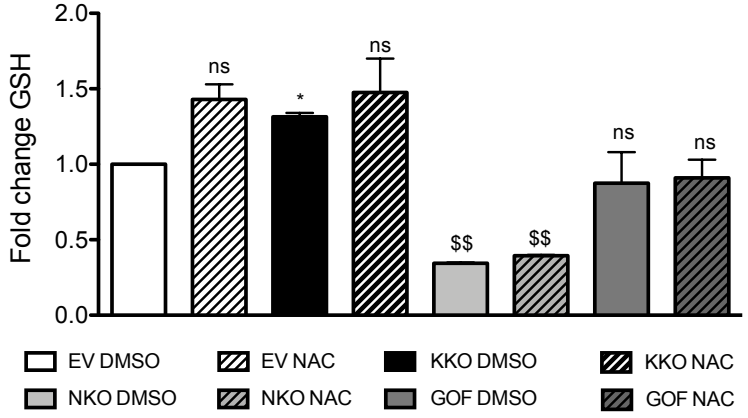

Fold change GSH
2.0
1.5
1.0
0.5
0.0
ns
*
ns
$$
$$
ns
ns
EV DMSO
EV NAC
KKO DMSO
KKO NAC
NKO DMSO
NKO NAC
GOF DMSO
GOF NAC

Figure 5.2.19: KKO H1299 cells contain higher levels of total glutathione than GOF H1299 cells

*EV H1299 cells are represented by the white bar, KKO H1299 cells are represented by the black bar, NKO H1299 cells by the light grey bar and GOF H1299 by the dark grey bar, on each of the three graphs. The data obtained after treatment with either BSO or NAC is shown in the striped bars. (A) Total glutathione levels were measured using the mCB fluorescent probe. Cells were seeded in 10% FCS supplemented DMEM and once a confluency of 90% was reached, plates were washed thoroughly and incubated in the dark in the presence of 20 μM mCB- HBSS for 30 min before reading fluorescence. (B) Total glutathione levels were measured as above after 100 μM BSO treatment for 24 hr. (C)Total glutathione levels were measured as above after 5 mM NAC treatment for 3 hr. The bars displayed each graph represents the mean value obtained from three independent biological replicates minus the control well and normalized total protein concentration, and the associated SEM. Statistically significant alterations relative to the EV H1299 or EV H1299 DMSO treated data are denoted on the graphs. All statistical analysis was carried out using GraphPad Prism 5 software. All significant increases are denoted as follows: *, P< 0.05; **, P<0.01; *** P< 0.001. P values > 0.05 were deemed not significant and are denoted by "ns". Significant decreases are denoted with a $ sign.*

5.2.21 The increase in total glutathione levels seen in the KKO H1299 cells may be due to increased reduced glutathione.

Glutathione exists in one of two states in the cell; either as GSH or GSSG. The GSH: GSSG ratio is used as an indicator of cell health. To determine the levels of GSH and GSSG in the cell, a assay adapted from Rahman et al. (2007) was performed. As shown is Figure 5.2.20, KKO H1299 cells contain higher total glutathione, GSH, and GSSG levels when compared to either the EV or GOF H1299 cell lines. Most noticeably, the levels of GSH appear to be very robustly elevated in the KKO cell line when compared to the GOF H1299 cells.

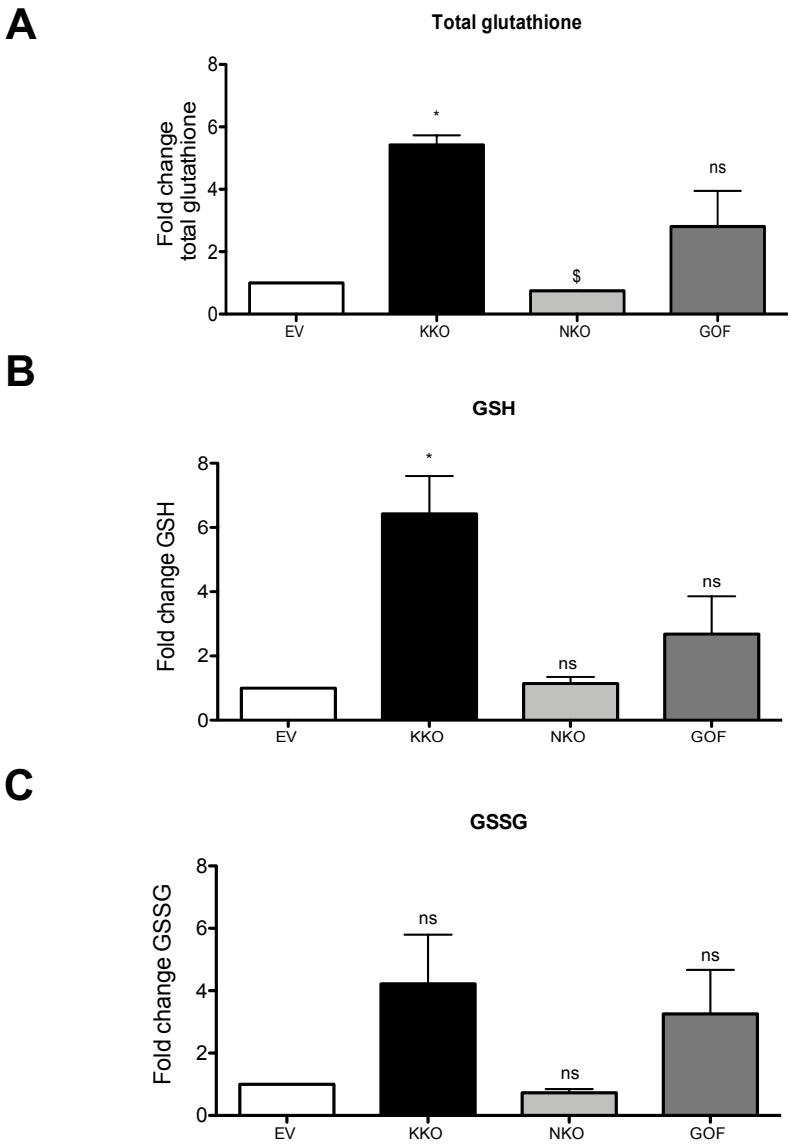


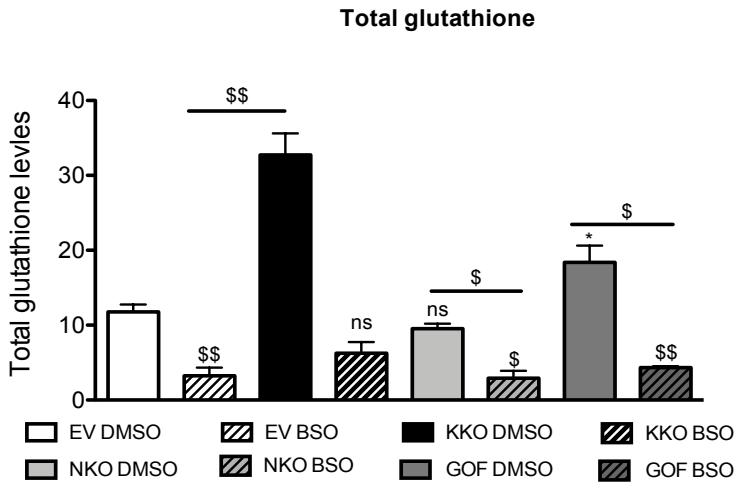
Figure 5.2.20: KKO H1299 cells contain higher levels of glutathione than GOF H1299 cells

*EV H1299 cells are represented by the white bar, KKO H1299 cells are represented by the black bar, NKO H1299 cells by the light grey bar and GOF H1299 by the dark grey bar, on each of the three graphs. (A) Analysis of total glutathione levels across the panel of H1299 CRISPR/Cas9 cell lines. (B) Analysis of GSH levels across the panel of H1299 CRISPR/Cas9 cell lines. (C) Analysis of GSSG levels across the panel of H1299 CRISPR/Cas9 cell lines. The bars displayed each graph represents the mean value obtained from three independent biological replicates minus the control well and normalized total protein concentration, and the associated SEM. Statistically significant alterations relative to the EV H1299 cell lines are denoted on the graph. All statistical analysis was carried out using GraphPad Prism 5 software. All significant increases are denoted as follows: *, P< 0.05; **, P<0.01; *** P< 0.001. P values > 0.05 were deemed not significant and are denoted by "ns". Significant decreases are denoted with a $ sign.*

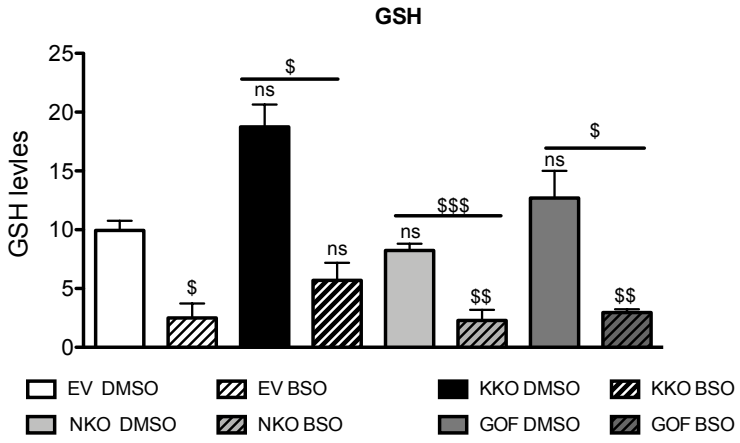
5.2.22 The GSH:GSSG ratio can be altered in both KKO and GOF H1299 cells through treatment with BSO.

To determine whether the levels of GSH and GSSG can be altered by treatment with BSO the same assay as described above in section 5.2.21 was carried out in the presence of BSO. The data in Figure 5.2.21, show that treatment with BSO leads to a significant decrease in GSH, GSSG and total glutathione levels in both the KKO and GOF cell lines, with the effect of BSO being more pronounced on GSSG levels than those of GSH. The largest decrease in total, reduced and oxidized glutathione levels was seen in KKO H1299 cells, suggesting that maintenance of their glutathione levels is more sensitive to inhibition of continuing biosynthesis than is the maintenance of glutathione levels in the other cell lines.

A



B



C

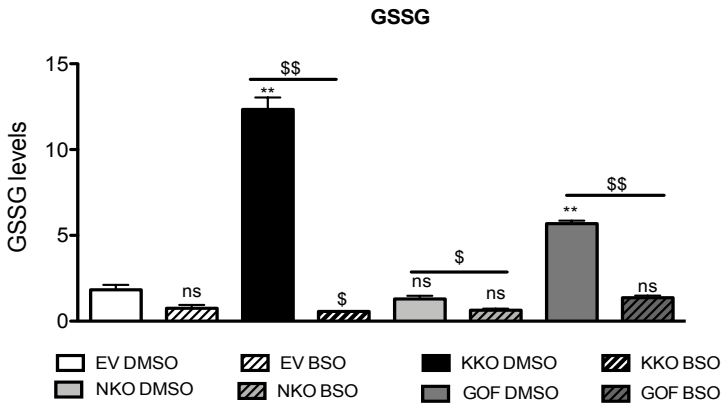


***Figure 5.2.21*: KKO H1299 cells contain greater levels of total glutathione than GOF H1299 cells due to augmented levels of GSH.**

*EV H1299 cells are represented by the white bar, KKO H1299 cells are represented by the black bar, NKO H1299 cells by the light grey bar and GOF H1299 by the dark grey bar, on each of the three graphs. the data obtained after treatment with BSO is shown in the striped bars. (A) Total glutathione levels were measured across the four CRISPR/Cas9 cell lines, in the presence or absence of BSO. Cells were treated with 100 μM BSO for 24 hr prior to analysis. (B) GSH levels were measured across the four CRISPR/Cas9 cell lines, in the presence or absence of inhibitor BSO. Cells were treated with 100 μM BSO for 24 hr prior to analysis (C) GSSG levels were measured across the four CRISPR/Cas9 cell lines, in the presence or absence of BSO. Cells were treated with 100 μM BSO for 24 hr prior to analysis. All treatments were carried out in DMEM supplemented with 10% FCS. The bars displayed each graph represents the mean value obtained from three independent biological replicates minus the control well and normalized total protein concentration, and the associated SEM. Statistically significant alterations relative to the DMSO treated EV H1299 cell line are denoted on the graph and significant alterations upon treatment between the different cell lines are denoted above the black line. All statistical analysis was carried out using GraphPad Prism 5 software. All significant increases are denoted as follows: *, P< 0.05; **, P<0.01; *** P< 0.001. P values > 0.05 were deemed not significant and are denoted by "ns". Significant decreases are denoted with a $ sign.*

5.2.23 KKO H1299 cells are more markedly re-sensitized to cisplatin chemotherapy by co-administration of BSO than GOF H1299 or EV H1299 cells.

In cancer cells, increased GSH synthesis has been associated with development of drug resistance through an increased capacity of GSH to form GSH-conjugates with antineoplastic drugs, which can lead to the more rapid export and elimination of anticancer drugs from cells. To determine whether treatment with BSO might result in a decrease in GSH levels that leads to increased sensitivity to cisplatin, cell viability was accessed using the MTT assay. To this end, EV, KKO, NKO and GOF H1299 cells were treated with a serial dilution of cisplatin, alone or in combination with 100 μM BSO, for 72 hr. Data in Figure 5.2.22 and Table 5.2.6 show that the most significant change in sensitivity to cisplatin treatment in the presence of BSO was in KKO H1299 cells. Although it should be taken into consideration that the IC50 values displayed will be estimated values extrapolated from the graphs given below. It should also be noted that surprisingly in this assay the GOF H1299 cells did not show an increased IC_{50} towards cisplatin alone when compared with that in EV H1299 cells. It would be assumed that the increased NRF2-signaling in the GOF cells would have made them more resistant to cisplatin as seen in the KKO cells. At present, it is difficult to draw conclusions about why GOF cells do not display increased resistance to cisplatin.

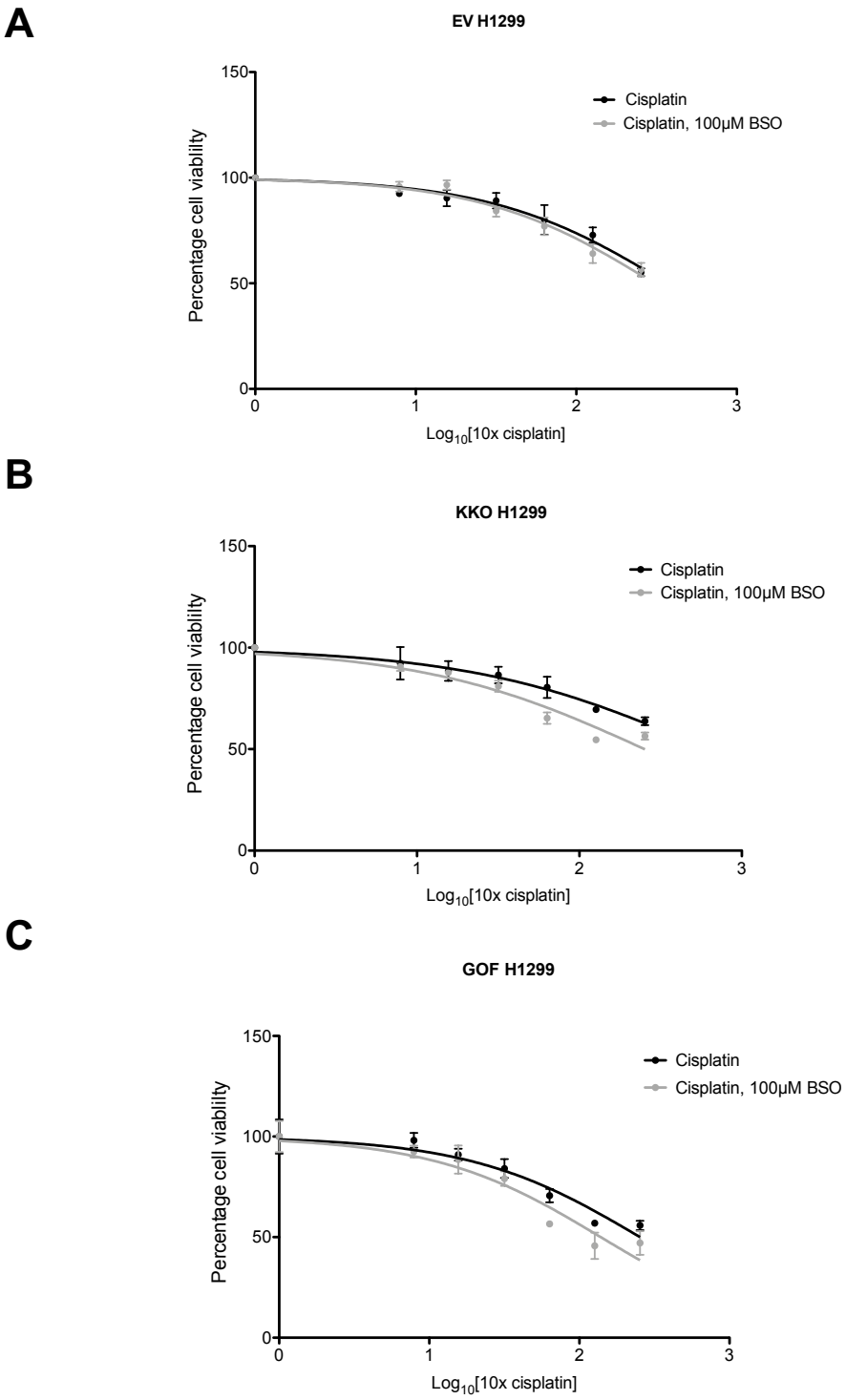


Figure 5.2.22: **KKO H1299 cells can be re-sensitized to cisplatin by co-administration of BSO.**

(A-C) The cell viability of the four CRISPR/Cas9 cell lines in the presence of either a range of cisplatin concentrations or cisplatin combined with 100 μM BSO, was accessed via MTT assay. The data displayed is one representative graph from three independent biological replicates with the associated percentage error. The bars displayed each graph represents the mean value obtained from three replicates minus the control well and the associated percentage error.

	Cell line							
		EV H1299		KKO H1299		GOF H1299		
		Cisplatin	Cisplatin, 100μM BSO	Cisplatin	Cisplatin, 100μM BSO	Cisplatin	Cisplatin, 100μM BSO	
	Average IC50 (μM)	30.37	24.01	57.27	18.07	28.60	18.77	
	StDev	5.65	5.86	2.82	6.19	3.83	5.16	
	Significance (Relative to the Cisplatin alone control for that cell line)	*		**		NS		

Table 5.2.6: Variation in cisplatin IC50, with and without co-administration of BSO, across the four CRISPR/Cas9 cell lines.

*Average value for three independent biological replicates is displayed in the table, along with the associated standard deviation (StDev). All statistical analysis was carried out using GraphPad Prism 5 software. All significant increases are denoted as follows: *, P< 0.05; **, P<0.01; *** P< 0.001. P values > 0.05 were deemed not significant and are denoted by "ns". Significant decreases are denoted with a $ sign.*

5.3 Discussion

In this chapter, cell lines with *KEAP1* mutations have been shown to exhibit a more pronounced "NRF2-gene expression signature" than cell lines with *NFE2L2* mutations. The enhanced NRF2-signalling seen in *KEAP1* mutant cell lines seems to result in them being more dependent on certain NRF2-regulated pathways, which can consequently be targeted therapeutically. Here it has been highlighted that cell lines that harbour a mutation in *KEAP1* are more dependent on glutathione biosynthesis and therefore more sensitive to glutathione inhibition.

5.3.1 The advantages of the CRISPR/Cas9 cell line approach to generate mutant lung cancer cell lines

To validate our bioinformatics findings, we initially used a panel of six commercially available lung cancer cell lines: amongst theses, two harboured *KEAP1* mutations, two harboured *NFE2L2* mutations and two expressed both wildtype *KEAP1* and *NFE2L2*. Characterisation of these cell lines showed similar elevations in mRNA and protein levels of NRF2-target genes in the *KEAP1* mutant cell lines when compared with the *NFE2L2* mutant cell lines. However, because of different genetic backgrounds and somatic mutations in these commercially-available cell lines, it was not possible to attribute such changes solely to the presence of mutant *KEAP1* or mutant *NFE2L2*.

In an attempt to analyse the effect of mutations in just *KEAP1* or *NFE2L2* on gene expression and phenotype, the isogenic H1299 CRISPR/Cas9 panel was generated. Nevertheless, this CRISPR/Cas9-generated panel is not without its limitations. In particular, part of the CRISPR/Cas9 methodology involves clonal selection, which introduces clone-specific variation. In this chapter, only one clone per CRISPR/Cas9 cell line has been studied. Thus, to robustly validate the findings several key experiments ought to be repeated with at least two other clones per cell line. Another possible weakness in the experimental design is that the cell lines generated either have a knockout of *KEAP1* or a partial deletion of the Neh2 domain of NRF2, and these do not accurately mimic the types of mutations found in lung cancer; for example, KEAP1 protein is never fully deleted in lung cancer cells, but rather is subject to single amino acid substitution

mutations. To generate a more human representative model, cell lines could be generated by CRISPR/Cas9 with single amino acid substitutions to mimic the exact mutations observed in human tumours. Also, mutations in *KEAP1* and *NFE2L2* do not occur alone in cancer. It is thought that mutations that dysregulate NRF2 occur to support the promotion phase of carcinogenesis but are not fundamental for cancer initiation, nor are the sufficient to initiate cancer. They therefore coexist with mutations in common oncogenic genes such as *KRAS* and *TP53*. In this context, it should be noted that the presence of *KEAP1* mutations alone have been reported to produce no adverse phenotype in ADC mouse models (Best et al., 2018). Also, we have generated preliminary data in the lab suggesting that alteration in KRAS signalling will affect NRF2 activation and NRF2-target gene expression (see, Appendix 8.9). Therefore, to get a better understanding of the role that *KEAP1* mutations and *NFE2L2* mutations have on NRF2-signalling during tumourigenesis it would be useful to make combinational CRISPR/Cas9 cell lines with alterations in several genes, mirroring what is observed in human tumours.

5.3.2 Why would KEAP1 mutants have more pronounced signalling?

From the data presented in this chapter using both the commercial panel of six cell lines and the isogenic H1299 CRISPR/Cas9 panel, it is clear that mutations in *KEAP1* and *NFE2L2* are different in terms of their effects on NRF2-signalling.

We have briefly eluded to a potential reason for this in Figure 5.2.18, were the protein expression of four KEAP1-binding proteins was analysed in the KKO and GOF H1299 cells. Ben Major and colleagues, carried out a comprehensive mass spectrometry-based analysis to identify proteins that possess an ETGE or ESGE and can interact with KEAP1 (Hast et al., 2013). It is thought that these KEAP1-interacting proteins can function as indirect regulators of NRF2 through outcompeting it for binding to KEAP1, which would lead to an increase in NRF2 activity. They identified 42 KEAP1-binding proteins, out of which 17 contained either an ETGE motif, an ESGE motif or both motifs. In this chapter we looked at the expression of only four of these 17 KEAP1-interacting proteins, namely: TSC22D4 (which contains an ETGE motif), NRF1 (which contains both an DLG and ETGE motif), p62/SQSTM1 (which contains a phosphorylatable STGE motif

and inhibits the interaction of the DLG domain of NRF2 with KEAP1 (Komatsu et al., 2010)) and PGAM5 (which has an ESGE motif). Surprisingly, the abundance of these four proteins was not significantly altered across our CRISPR/Cas9 panel of H1299 cell lines, particularly in the KKO cells, and therefore seem unlikely to account for their altered NRF2-target gene expression. This suggests that potentially the expression of KEAP1-interacting proteins is not altered in this panel. However, it is more likely that just the four proteins we chose to analyse were not altered. Moving forward, it would be interesting to look at the expression of other KEAP1-interacting proteins such as dipeptidyl peptidase 3 (DPP3) which has been shown to be elevated in cancer and interact with NRF2 (Lu et al., 2016).

5.3.3 Mutations in the ETGE region of NRF2 do not always lead to an elevation in NRF2-signaling

Mutations in *NFE2L2* in lung cancer have been shown to occur in the region of the gene encoding the Neh2 domain, with all mutations clustering to amino acids 24-34 and 75-82, in the Neh2 region (Hayes and McMahon, 2009; Fukutomi et al., 2014). This strict focal clustering of mutations is a common feature of oncogenic genes. Mutations in the two regions of *NFE2L2* that are closely associated with the DLG and ETGE motifs have been shown in impair KEAP1-mediated degradation of NRF2 and result in increased transactivation by the transcription factor. This is due to the Neh2 domain not requiring to undergo a post-translational modification event before KEAP1 is able to direct NRF2 for proteasomal degradation, and so suggests the structural integrity of the DLG and ETGE motifs is critical for the interaction between NRF2 and KEAP1 (Shibata et al., 2008).

In an initial attempt to validate the bioinformatics we used a commercial panel of six cell lines, of which H2228 and H1568 harboured mutations in *NFE2L2*. These cell lines were chosen as they expressed mutations in either the DLG or ETGE motifs of NRF2, which would suggest that they have high NRF2-signalling. However, when the expression of NRF2-target genes in these two cell lines was analysed (shown in Figure 5.2.5 and 5.2.6) H2228 cells appeared to have the expected elevation in NRF2-target gene expression but H1568 cells showed the opposite. This was surprising for two reasons: (1) mutations in the DLG and ETGE motifs have often shown to be similar in terms of their effect in

NRF2 signalling; (2) H1568 cells have a mutation in the high affinity ETGE motif and H2228 cells possess a mutation in the low affinity DLG motif. Work by (Shibata et al., 2008) demonstrated that lung cancer cell lines with mutations in the DLG motif can still be partially bound by KEAP1, probably due to the presence of an intact ETGE high affinity site. Therefore, out of our two chosen *NFE2L2* mutant cell lines, the H1568 cells with a mutated ETGE and intact DLG motif would be expected to show higher NRF2-target gene expression than the H2228 cells, with a mutated DLG and an interact ETGE motif, but we see the opposite. Potentially, this could be due to these two cell lines harbouring mutations in other genes that are affecting NRF2-target gene expression. Both of these cell lines also harbour homozygous mutations in TP53, which are thought to be oncogenic. No other mutations *in KEAP1, KRAS, CUL3, PIK3CA, PTEN, STK11* or *EGFR* were found in H2228 or H1568 cell lines. Potentially one of these cell lines could possess a mutation in gene that is dramatically influencing NRF2-signalling in a manner that we have not explored.

At this point we are unable to account for the surprising differences in NRF2-target gene expression between these two *NFE2L2*-mutant cell lines, this would require a more in-depth analysis of the KEAP1 binding ability in a considerably larger panel of *NFE2L2* mutant cell lines.

5.3.4 Mutations in *KEAP1* and *NFE2L2* alone are not sufficient to induce the expression of NRF2-target genes involved in the pentose phosphate pathway (PPP).

Masayuki Yamamoto and Hozumi Motohashi and their colleagues published an in-depth analysis into the role that NRF2-signalling plays in metabolism. In particular, they demonstrated that NRF2 has the ability to activate several metabolism genes, including but not limited to, glucose-6-phosphate dehydrogenase (G6PD) and phosphogluconate dehydrogenase (PGD) (Mitsuishi et al., 2012). In Figure 5.2.14 we accessed the protein expression of both of these metabolism-associated genes across our panel of H1299 CRISPR/Cas9 cell lines and found surprisingly that the protein expression of both G6PD and PGD was unaltered in the four cell lines and unaffected by CDDO-ME treatment. There are several factors that may account for this unusual result. Firstly; we could have been “unlucky” in our choice of metabolism markers, in that G6PD and PGD may

not be altered but other metabolism genes could be. To get a better understanding we could look at the expression of *TKT, TALDO1* and *PPAT*, which were also shown to be upregulated by NRF2. Secondly, in the fore mentioned publication by the Yamamoto group they analysed the expression of metabolism genes in cell lines harbouring mutations in *KEAP1*, not the expression of these genes in cell lines with *NFE2L2* mutations. This would not account for why we failed to see a change in the protein levels of G6PD or PGD in the KKO H1299 cells but potentially could explain why we did not see expression changes in the GOF H1299 cells. The third, and most likely, explanation for this data is that our CRISPR/Cas9 panel do not harbour any mutations that would lead to a disruption of PTEN-PI3K–AKT signalling. As highlighted by Kate Sutherland and colleagues, only when PTEN signalling was abolished to create aberrant AKT signalling in the absence of KEAP1, was an increased metabolic signature seen. Importantly this metabolic activation was not seen in the presence of mutations in *KEAP1* alone (Best et al., 2018). Since we do not have elevated AKT signalling in the H1299 cell lines, we may not see changes in the expression of metabolism associated proteins. It would be interesting to siRNA knock out *PTEN* in the KKO H1299 cells and see if this leads to increased expression of NRF2-target genes associated with the PPP. However, care should be taken when combining mutations in *PTEN* and *KEAP1* to study metabolism, as these two mutations are mutually exclusive, at least in lung cancer (this concept is discussed in more detail in Chapter 4).

5.3.5 Are differences in the zygosity of *KEAP1* and *NFE2L2* mutant H1299 cell lines responsible for their distinct phenotypes?

It became evident from the bioinformatics analyses (discussed in full in Chapter 4) that the majority of *KEAP1* mutations are homozygous whereas *NFE2L2* mutations are heterozygous. To determine whether the enhanced NRF2-gene expression signature that we see in *KEAP1*-mutant H1299 cells is not because we are comparing a homozygous mutation with a homozygous mutant *NFE2L2* mutant cell line, we sent our cell lines to undergo shotgun cloning followed by sequencing. The sequencing results (Appendix 8.10) show that the KKO H1299 cell line that we have been using in this chapter is a KEAP1 homozygous KO, whereas the GOF H1299 cell line is heterozygous for

a deletion of amino acids 28-35 in the Neh2 domain of NRF2. This means that the two cell lines that we have generated do veraciously mimic human lung cancer cell lines. Although it would be extremely interesting to generate a homozygous *NFE2L2* mutant cell line to see whether it would have a similar elevated NRF2-gene expression signature as a homozygous *KEAP1* mutant cell line, such mutations would appear to be rare in cancer cells. It might therefore be more appropriate to select a *KEAP1* heterozygous KO H1299 clone from amongst our KKO pool of cells and compare them with GOF H1299 cells.

It is also worth noting that sequencing revealed that disruption of the Neh2 domain within the *NFE2L2* gene means that isoform 6 of NRF2 is still expressed in these cell lines. Isoform 6 is minor isoform composed of 505 amino acids which contains a CNC-bZIP domain. This may in part account for the partial detection of NRF2 protein expression in the NKO H1299 cell line, but this would require further investigation.

5.3.6 Concluding remarks

Mutations in *KEAP1* and *NFE2L2* have always been grouped together and treated as being equally effective in terms of their ability to upregulate NRF2-directed ARE-driven gene expression. Also, the majority of papers that are concerned with NRF2 signalling tend to utilise *KEAP1*-mutant cells and cells harbouring *NFE2L2* mutants are rarely used. In this chapter, we have attempted to demonstrate through the use of bioinformatics, commercial cell lines and isogenic CRISPR/Cas9 cell lines, that *KEAP1* and *NFE2L2* mutations are in fact different. We have demonstrated that *KEAP1* mutant cells have higher NRF2-signaling in comparison to *NFE2L2* mutant cell lines. In particular, *KEAP1*-mutant cells have enhanced glutathione biosynthesis and seem to be dependent on this pathway for survival because they are more susceptible to inhibition by BSO. This presents the idea that through targeting specific NRF2-regulated pathways in tumours we can enhanced the effect of currently available chemotherapeutic agents. However, to fully validate this model we would have to expand our panel of commercial cell lines and generate CRISPR/Cas9 cell lines that mimic human cancer better. The data presented in this chapter is just the start of what could be a very exciting discovery.

6.0 Final discussion and future perspectives

Cells are constantly battling against the damaging effects ROS and are therefore relent on the activity of the redox-regulated transcription factor, NRF2 to orchestrate protection against such a threat. Since the discovery of the role of NRF2 as the master regulator of the cell oxidative stress responses, the field of NRF2 research has expanded exponentially. This is in part due to the identification of over 200 ARE-containing genes that are trans-activated by NRF2, in response to a variety of stressors. NRF2 is now accepted to be critical for several cell pathways such as: metabolism, detoxification, antioxidant production, purine synthesis, drug efflux, iron-binding, NADPH generation, proteasome subunit assembly, apoptosis and autophagy; the majority of which can be regarded as cytoprotective (Hayes and Dinkova-Kostova, 2014; Tebay et al., 2015).

Aberrant expression of NRF2 is commonly associated with lung cancer and was initially implicated in causing the increased expression of cytoprotective enzymes that were found in resistant NSCLC cells (Homma et al., 2009). However, more recently it has been shown that NRF2 can play a role in development of each of the hallmarks of cancer including, metabolic reprograming leading to increase glycolysis, continuous proliferative signalling, decreased apoptotic cell death, increased replicative capacity, enhanced angiogenesis, prototoxic stress, high levels of genomic instability, increased tissues invasiveness and metastatic potential, ability to avoid immune defences, resistance to antigrowth signals and increased tumorigenic inflammatory signalling (Rojo de la Vega et al., 2018; Hanahan and Weinberg, 2000; Hanahan and Weinberg, 2011).

Pervious work has shown that mutations in *NFE2L2, KEAP1* and *CUL3* occur in lung cancer and result in increased NRF2 activity through inhibition of KEAP1-directed degradation of the transcription factor. In this thesis, we have investigated mutations in *KEAP1, NFE2L2* and *CUL3* separately to access their individual impact on NRF2 activity using NRF2-target gene expression as a read out and studied the influence that a priming kinase plays in KEAP1-independent degradation of NRF2. Through this, we hoped to gain a greater insight into the complex role that NRF2 plays in oncogenesis and the potential of stimulating

alternative NRF2 degradation pathways to decrease the expression of the transcription factor in lung cancer tumours harbouring the aforementioned mutations.

6.1 NRF2 inhibitors – history and limitations

Hyper-activation of NRF2 signalling is seen in several forms of cancer and correlates with poor prognosis and enhanced resistance to platinum-based antineoplastic drugs and nitrogen mustards. Therefore, in the field there has been a momentous drive to produce specific and effective NRF2 inhibitors, to no avail.

One of the most promising candidates was the natural derived quassinoid compound, brusatol. The antitumor effects of brusatol were documented in the 1970s (Hall et al., 1979; Elgebaly et al., 1979) and again in the early 2000s (Mata-Greenwood et al., 2002) but it was not until 2011 that work published by Donna Zhang's group identified this compound as a reversible inhibitor of NRF2. Brusatol has been shown in both *in vitro* and *in vivo* models to ameliorate chemoresistance to cisplatin through inhibiting NRF2 protein expression (Ren et al., 2011). Subsequent research carried out by other groups has now demonstrated that nanomolar concentrations of brusatol leads to the depletion of NRF2 protein expression independently of KEAP1, the UPS and any currently documented form of NRF2 regulation, suggesting it is through a currently uncharacterised mechanism (Olayanju et al., 2015). However, even though this compound is specific for NRF2, as it has been shown to not affect the protein levels of several other proteins with both short and long half-lives (Olayanju et al., 2015), but it is not specific for cancer cells and has the potential to sensitize non-cancer cells to the damaging effects of antineoplastic drugs. Although it should be noted that other groups have published contrasting data showing that brusatol inhibits RNA, DNA and global protein synthesis in the cell (Hall et al., 1979). Global inhibition of NRF2 in all cell types through the administration of brusatol would not be advantageous as NRF2 is responsible for cell defences that are fundamental for normal cell survival. However, due to cancer cells having enhanced levels of ROS they maybe more sensitive to NRF2 inhibition than normal cells, and this could be potentially exploited by using the correct concentration of brusatol. Due to the lack of a characterized mechanism of action

and the potential non-target effects, a lot of work will still need to be carried out before brusatol reaches clinical administration.

ML385 is a thiazole-indoline compound and another potentially promising NRF2 inhibitor (Singh et al., 2016). ML385 directly interacts with the Neh1 domain of NRF2, prevents dimerization and at micromolar concentrations has been shown in *in vivo* NSCLC models to re-sensitize cells to carboplatin leading to reduced tumour proliferation. One limitation of ML385 is that it may potentially target other transcription factors, aside from NRF2 (Singh et al., 2016).

6.2 KEAP1-independent repression of NRF2 is influenced by a priming kinase

An alternative non-pharmacological mechanism by which the activity of NRF2 can be decreased, is to stimulate the $SCF^{\beta\text{-}TrCP}$ mediated degradation pathway. This is a particularly attractive concept as it would potentially ameliorate the effects of NRF2 upregulation in tumours harbouring mutations in *KEAP1, NFE2L2* and *CUL3*. Research into KEAP1-independent mechanisms of NRF2 degradation has peaked in recent years, with it now being widely accepted that GSK-3 mediates phosphorylation of the Neh6 domain. This is fundamental for the formation of the phosphodegron to which the substrate adaptor protein β-TrCP binds leading to the ubiquitination of NRF2 and subsequent 26S proteasomal degradation of the transcription factor. In Chapter 3 of this thesis, we have continued to investigate this form of NRF2 regulation, focusing on the notion that a priming kinase may work in concert with GSK-3. We have demonstrated, through using the DYRK family of kinases as an example, that phosphorylation of Ser-347 functions as a priming event leading to subsequently enhanced phosphorylation by GSK-3 at Ser-342 and Ser-338, accentuating $SCF^{\beta\text{-}TrCP}$ mediated degradation. We have also shown that regulation of this priming event would affect cell proliferation and chemosensitvity to platinum-based therapies, through altering GSK-3 signalling without directly targeting GSK-3 itself. However, development of specific small molecule activators of the priming kinase or peptides that would mimic the priming phosphate and bind to GSK-3 locking it in a “primed state”, is required.

We have also indicated that we believe that another kinase mediates the phosphorylation of Ser-335. Since β-TrCP binding to a substrate is enhanced by

the presence of two phosphorylated neighbouring residues, targeting the kinase that phosphorylates Ser-335 would provide another mechanism by which to regulate NRF2 degradation independently from KEAP1 and another novel potential therapeutic target. Currently we do not know the identity of the kinase that may target Ser-335 but as mentioned in the discussion of Chapter 3, we believe that due to previous literature linking CK1, DYRK and GSK-3 together, unpublished data from our lab identifying FXXXF docking motifs in NRF2 and also the extensive homology of the amino acids sequence surrounding the 335 site with that of a CK1 consensus sequence, that potentially CK1 mediates the phosphorylation at this site.

6.3 Bioinformatics analysis of the prevalence, co-occurrence and impact of mutations in *KEAP1* and *NFE2L2* in lung cancer cell lines and tumours.

Mouse models are a valuable tool for analysing the organism-wide effect of alterations in the expression of a particular gene. NRF2 KO mice are viable with normal embryonic development and no outwardly adverse phenotype (Chan et al., 1996) . However, through carrying out ageing studies with these animals and exposing them to various stimuli, it has been revealed that NRF2 KO mice have impaired liver regeneration (Beyer et al., 2008), are at more risk of developing cancer (Khor et al., 2008; auf dem Keller et al., 2006), are more susceptible to toxins (Enomoto et al., 2001), develop retinopathies (Zhao et al., 2011), have impaired immune systems (Jiang et al., 2004; Innamorato et al., 2008) and defective cardiac function (Erkens et al., 2015).

Hyper-activation NRF2 has also been studied in mice, through the disruption of *KEAP1* but rarely through a gain-of-function mutation in the *NFE2L2* gene. One group has studied the effect of NRF2 hyper-activation in human patients that have extensive multisystem disorder due to *de novo* mutations in *NFE2L2.* The four patients studied have mutations in the Neh2 domain that have previously been described in cancer cells and present with various symptoms, such as: immunodeficiency's, neurological symptoms, and hypohomocysteinaemia (Huppke et al., 2017). Murine models with altered NRF2 levels have been a particularly valuable tool in identifying novel NRF2-target genes that play a role in carcinogenesis. However, often a disruption in *KEAP1* is paired with several disruptions in other oncogenic genes. This is

understandable as it will lead to cancer, permitting the study role of NRF2 in the progression/severity of the disease, and often a combination of alterations in several genes must happen to allow quicker initiation of the disease due to time constraints. Although caution should be taken to make sure that the cancer model in mice is a true reflection of the human disease and does not have genes altered that are not seen altered alongside *KEAP1* mutations in human disease. In this thesis, we have shown that *KEAP1* and *NFE2L2* mutations should be treated separately. Thus, whilst *KEAP1* and *NFE2L2* mutations do not co-exist in either lung cancer tumours or cell lines, each co-exists alongside different mutations and they also have different effects on NRF2-target gene expression. One of the key findings of Chapter 4 was that *KEAP1* mutations co-exists with mutations in *KRAS* and are mutually exclusive to mutations in *NFE2L2* and *TP53*. Whereas, mutations in *NFE2L2* co-occur with mutations in *TP53* and are mutually exclusive from mutations in *KEAP1* and *KRAS.* Often in mouse lung tumourigenesis models that relate to NRF2, *KEAP1* mutations are combined with *PTEN* and *KRAS* mutations, a combination that we believe does not naturally occur in human lung cancer.

Another key finding of Chapter 4 was that mutations in *KEAP1* and *NFE2L2* show different zygosities, with mutations in *KEAP1* being predominantly homozygous and mutations in *NFE2L2* being predominantly heterozygous. This potentially accounts for the data we obtained using commercially available cell lines and our generated H1299 CRISPR-Cas9 panel, showing that an alteration in *KEAP1* has a more pronounced elevation in NRF2-target genes expression, in contrast to a *NFE2L2* alteration. As heterozygous *NFE2L2* mutant cell lines will still poses one allele encoding wildtype NRF2, a portion of the transcription factor in the cell is still subject to degradation by KEAP1. By contrast, a homozygous disruption in *KEAP1* will mean that these cell lines will be completely devoid of KEAP1-mediated NRF2 degradation. The reason behind the variation in zygosities between *KEAP1* and *NFE2L2* mutants in lung cancer is currently unknown. Generally, homozygous mutations are often associated with tumour suppressor genes and mutations that lead to a loss of function. It is thought that mutations in tumour suppressor genes are “molecularly recessive” (Futreal et al., 2004) and therefore require both copies of the gene to be mutated to exert an effect. This in combination with the notion that mutations in tumour suppressor

genes tend to be spread throughout the protein showing no clear focal point, suggests that *KEAP1* may be a tumour suppressor gene (Baeissa et al., 2016). It has been extensively debated to whether *NFE2L2* can be classed as an oncogene with the field deciding that such a narrow classification would not be appropriate. However, the data generated in Chapter 4 of this thesis indicates that *NFE2L2* has a mutational pattern that is comparable to that of an oncogene with heterozygous mutations that all tend to occur in a particular region of the protein. Mutations in oncogenes are thought to be "molecularly dominate" and only one altered copy of the gene is required to cause an oncogenic phenotype, as seen in BRAF mutations (Baeissa et al., 2016). If we apply Vogelstein's 20:20 rule, which states that if 20% of mutations observed in a gene are missense mutations that occur at a single position in the sequence, *NFE2L2* would be classified as a oncogene (Vogelstein et al., 2013).

To gather a greater understanding of the impact of mutations in *KEAP1* and *NFE2L2,* we would suggest that the next logical step would be to generate *NFE2L2* homozygous gain-of-function H1299 cells using CRISPR-Cas9 gene editing, which could be compared directly to our already existing homozygous KEAP1 KO H1299 cells. Once we have carried out the required cell experiments, the next stage would be to combine this zygosity data with our data showing the co-occurrence of *KEAP1* and *NFE2L2* with other oncogenic mutations to generate a more robust and translationally relevant mouse model of lung carcinogenesis to study the role of NRF2 signalling.

6.4 *KEAP1* mutant cell lines increase glutathione biosynthesis more than *NFE2L2* mutant cell lines and are therefore more sensitive to inhibition of this pathway.

A great controversy centres around the role NRF2 plays in cancer. Several papers have documented the beneficial role that NRF2 plays in halting the progression of cancer in the early stages, but a wealth of papers have also highlighted that cancer cells increase NRF2 signalling to permit their survival. For example, the expression of the AKR family of proteins are robustly elevated upon increased NRF2 activity. The AKRs are well known for their detoxification potential but also metabolize carcinogens found in cigarette smoke and potentially enhance mutagenesis driving cancer progression (Penning, 2017).

This raises the argument of whether it would be safe to inhibit NRF2 therapeutically for the treatment of cancer. We have discussed at length in this thesis, the oncogenic effects of high levels of NRF2 in cancer due to somatic mutations in *KEAP1, NFE2L2 or CUL3* and the associated resistance to antineoplastic drugs that is seen in cell lines and tumours with aberrant NRF2 signalling. However, it is important to highlight the beneficial role that NRF2 plays in cancer prevention and the issue this raises in developing drugs to target the transcription factor.

Firstly, NRF2 KO mice are more susceptible to carcinogenesis than their wildtype counterparts (Khor et al., 2008). Also, several pharmacological activators of NRF2 have been shown to delay cancer progression, such as Sulforaphane and CDDO-ME (Beaver et al., 2018; Kim et al., 2012). SFN is a naturally occurring potent isothiocyanate compound found in cruciferous vegetables that activates NRF2 through impairing KEAP1 function, and has been shown in both *in vitro* and *in vivo* experiments to halt cancer progression (Zhang et al., 1994). KEAP1 contains several cysteine residues that function as thiol based cellular switches to regulate the activity of KEAP1, and subsequently NRF2, upon exposure to oxidative of electrophilic stress (Sporn and Liby, 2012). There are 27 cysteine residues in human KEAP1 that when modified alone, or in combination, lead to a conformational change in KEAP1 resulting in impaired KEAP1-mediated degradation of NRF2. SFN targets Cys-151 and at least three other Cys residues forming reversible thiol adducts, impairing the interaction between KEAP1 and CUL3 and leading to reduced ubiquitination of NRF2 (Hu et al., 2011a). SFN has been shown to inhibit the progression of carcinogenesis in the early stages of skin, lung, breast, colon and stomach cancer. However, the range of proteins effected by SFN administration is dose dependent. At low nanomolar concentrations of SFN, KEAP1 is inhibited but at higher concentrations SFN inhibits other proteins such as NF-κB and MAPK (Sporn and Liby, 2012; Kensler et al., 2013; Clarke et al., 2008).

CDDO methyl ester (CDDO-ME; also known as bardoloxone methyl or RTA 402) is a synthetic triterpenoid derived from Oleanolic acid. CDDO- ME, like SFN, forms a thiol adduct on Cys-151 of KEAP1, which increases NRF2 activity through inhibiting CUL3 binding to KEAP1 (Cleasby et al., 2014). Also like SFN, the number of CDDO-ME target proteins increases in a dose dependant manner,

and at low concentrations inhibits KEAP1 but at high millimolar doses also inhibits IKKβ and NF-κB (Ahmad et al., 2006). The antineoplastic effects of CDDO-ME have been extensively studied in both *in vitro* and *in vivo* models and has shown promising results in phase I clinical trials for advanced solid tumours and lymphoma (Hong et al., 2012). CDDO-ME can also function as a radioprotector, through neutralising the damaging effects of ROS due to radiotherapy in non-cancerous lung and breast cells (El-Ashmawy et al., 2014).

As NRF2 appears to play both a beneficial and detrimental role in cancer depending on the stage of the cancer and the smoking status of the patient, and the lack of both specific NRF2 activators or inhibitors, maybe targeting NRF2 itself is not the best therapeutic strategy. In Chapter 5 of this thesis we have demonstrated that inhibition of GSH biosynthesis has a more pronounced effect in *KEAP1* mutant than *NFE2L2* mutant cell lines and can re-sensitise *KEAP1* mutant cells to cisplatin. Therefore, potentially targeting the subgroup of NRF2-target genes that are upregulated is a better way of decreasing the cancer protective effects of NRF2 signalling in established later stage cancer. Glutathione biosynthesis is only one of the key NRF2 regulated pathways that is involved in carcinogenesis and potentially targeting other subgroups of NRF2-target genes e.g. drug detoxification genes will also show promising results.

6.5 Thesis summary and concluding remarks

The involvement of the transcription factor NRF2 in cancer has been studied extensively over the past decade. The field has made many rapid advancements in understanding the role of NRF2 in both cancer initiation and progression. However, a NRF2 biomarker has not been established. Biomarkers are cellular alterations that can be measured in patients tissues, cells or fluids, that can be used to study disease progression, diagnose disease, screen populations and measure the effect of treatments (Mayeux, 2004). *EGFR* and *KRAS* are two of the most commonly mutated genes in NSCLC and are associated with a negative disease outcome. The mutational status of *KRAS* and *EGFR* is often screened in lung cancer patients and is used as a predicative biomarker of whether patients will respond to EGFR inhibitors (Martin et al., 2013; Vincent et al., 2012). Since the expression of NRF2 itself is not a good read out for NRF2 activity, several NRF2 target genes have been suggested as

biomarkers of NRF2 activity such as AKR1B and 1C family members and NQO1. AKR1B10 expression is dramatically elevated in lung cancer tissues in comparison to neighbouring non-cancerous tissues, where it is undetectable (Zu et al., 2009). Also NQO1 expression has been shown to function as prognostic biomarker in pancreatic (Awadallah et al., 2008) and ovarian cancer (Cui et al., 2015). The establishment of a NRF2 biomarker would provide a greater insight into the fluctuations in NRF2 activity during disease and would dramatically aid in the development of specific NRF2 targeting compounds.

Overall in this thesis we have demonstrated the importance of a novel priming phosphorylation site within the Neh6 domain of NRF2 for $SCF^{\beta\text{-}TrCP}$ mediated degradation of the transcription factor and highlighted the differences between *KEAP1* and *NFE2L2* mutations in lung cancer.

ROS are often referred to as the “achilleas heel” of cancer cells due to the cancers high dependency on ROS for survival (Cui, 2012). However, oxidative stress is a hallmark of several diseases and degenerative disorders, not just cancer, highlighting the large therapeutic potential of targeting NRF2 signalling and the importance in studying the pathways/ events that lead to the dysregulation of NRF2 degradation.

7.0 References

Adler, V., Yin, Z., Fuchs Serge, Y., Benezra, M., Rosario, L., Tew Kenneth, D., Pincus Matthew, R., Sardana, M., Henderson Colin, J., Wolf, C. R., Davis Roger, J. & Ronai, Z. e. 1999. Regulation of jnk signaling by gstp. *The EMBO Journal,* 18**,** 1321-1334.

Ahmad, R., Raina, D., Meyer, C., Kharbanda, S. & Kufe, D. 2006. Triterpenoid cddo-me blocks the nf-κb pathway by direct inhibition of ikkβ on cys-179. *The Journal of Biological Chemistry,* 281**,** 35764-35769.

Ahrendt, S. A., Decker, P. A., Alawi, E. A., Zhu, Y.-r., Sanchez-Cespedes, M., Yang, S. C., Haasler, G. B., Kajdacsy-Balla, A., Demeure, M. J. & Sidransky, D. 2001. Cigarette smoking is strongly associated with mutation of the k-ras gene in patients with primary adenocarcinoma of the lung. *Cancer,* 92**,** 1525-1530.

Akerboom, T. P. M., Bilzer, M. & Sies, H. 1982. The relationship of biliary glutathione disulfide efflux and intracellular glutathione disulfide content in perfused rat liver. *The Journal of biological chemistry,* 257**,** 4248-52.

An, J. H., Vranas, K., Lucke, M., Inoue, H., Hisamoto, N., Matsumoto, K. & Blackwell, T. K. 2005. Regulation of the caenorhabditis elegans oxidative stress defense protein skn-1 by glycogen synthase kinase-3. *Proceedings of the National Academy of Sciences of the United States of America,* 102**,** 16275-16280.

auf dem Keller, U., Huber, M., Beyer, T. A., Kümin, A., Siemes, C., Braun, S., Bugnon, P., Mitropoulos, V., Johnson, D. A., Johnson, J. A., Hohl, D. & Werner, S. 2006. Nrf transcription factors in keratinocytes are essential for skin tumor prevention but not for wound healing. *Molecular and cellular biology,* 26**,** 3773-3784.

Awadallah, N., Dehn, D., J Shah, R., Russell Nash, S., K Chen, Y., Ross, D., Bentz, J. & Shroyer, K. 2008. Nqo1 expression in pancreatic cancer and its potential use as a biomarker. *Applied Immunohistochemistry and Molecular Morphology,* 16**,** 24-31.

Azadmanesh, J. & Borgstahl, G. 2018. A review of the catalytic mechanism of human manganese superoxide dismutase. *Antioxidants (Basel),* 7**,** 25.

Baba, K. & Miyazaki, T. 2016. Novel function of e3 ubiquitin ligase siah2 to regulate ros metabolism: Baba k et al . Novel function of siah2 in ros metabolism. *Journal of Biochemistry and Molecular Biology Research,* 2**,** 152-156.

Baba, K., Morimoto, H. & Imaoka, S. 2013. Seven in absentia homolog 2 (siah2) protein is a regulator of nf-e2-related factor 2 (nrf2). *The Journal of Biological Chemistry,* 288**,** 18393-18405.

Baeissa, Hanadi M., Benstead-Hume, G., Richardson, Christopher J. & Pearl, Frances M. G. 2016. Mutational patterns in oncogenes and tumour suppressors. *Biochemical Society Transactions,* 44**,** 925-931.

Baggstrom, M. Q., Stinchcombe, T. E., Fried, D. B., Poole, C., Hensing, T. A. & Socinski, M. A. 2007. Third-generation chemotherapy agents in the treatment of advanced non-small cell lung cancer: A meta-analysis. *Journal of Thoracic Oncology,* 2**,** 845-853.

Bain, J., Plater, L., Elliott, M., Shpiro, N., Hastie, C J., McLauchlan, H., Klevernic, I., Arthur, J Simon C., Alessi, Dario R. & Cohen, P. 2007. The selectivity of protein kinase inhibitors: A further update. *The Biochemical Journal,* 408**,** 297-315.

Bakhmutova-Albert, E., Yao, H., Denevan, D. & Richardson, D. 2010. Kinetics and mechanism of peroxymonocarbonate formation. *Inorganic Chemistry,* 49**,** 11287-96.

Banning, A., Deubel, S., Kluth, D., Zhou, Z. & Brigelius-Flohé, R. 2005. The gi-gpx gene is a target for nrf2. *Molecular and cellular biology,* 25**,** 4914-23.

Barja, G. 1999. Mitochondrial oxygen radical generation and leak: Sites of production in states 4 and 3, organ specificity, and relation to aging and longevity. *Journal of Bioenergetics and Biomembranes,* 31**,** 347-66.

Barnum, K. J. & O'Connell, M. J. 2014. Cell cycle regulation by checkpoints. *Methods in molecular biology (Clifton, N.J.),* 1170**,** 29-40.

Barone, M. C., Sykiotis, G. P. & Bohmann, D. 2011. Genetic activation of nrf2 signaling is sufficient to ameliorate neurodegenerative phenotypes in a drosophila model of parkinson's disease. *Disease Models & Mechanisms,* 4**,** 701-707.

Beaver, L. M., Löhr, C. V., Clarke, J. D., Glasser, S. T., Watson, G. W., Wong, C. P., Zhang, Z., Williams, D. E., Dashwood, R. H., Shannon, J., Thuillier, P. & Ho, E. 2018. Broccoli sprouts delay prostate cancer formation and decrease prostate cancer severity with a

concurrent decrease in hdac3 protein expression in transgenic adenocarcinoma of the mouse prostate (tramp) mice. *Current Developments in Nutrition,* 2**,** 3.

Berberat, P. O., Dambrauskas, Z., Gulbinas, A., Giese, T., Giese, N., Künzli, B., Autschbach, F., Meuer, S., Büchler, M. W. & Friess, H. 2005. Inhibition of heme oxygenase-1 increases responsiveness of pancreatic cancer cells to anticancer treatment. *Clinical Cancer Research,* 11**,** 3790-3798.

Berg , J., Tymoczko, J. & L., S. 2002. Chapter 18, oxidative phosphorylation. *In:* FREEMAN, W. (ed.) *Biochemistry 5th edition.* New York

Best, S. A., De Souza, D. P., Kersbergen, A., Policheni, A. N., Dayalan, S., Tull, D., Rathi, V., Gray, D. H., Ritchie, M. E., McConville, M. J. & Sutherland, K. D. 2018. Synergy between the keap1/nrf2 and pi3k pathways drives non-small-cell lung cancer with an altered immune microenvironment. *Cell Metabolism,* 27**,** 935-943.

Beyer, T. A., Xu, W., Teupser, D., auf dem Keller, U., Bugnon, P., Hildt, E., Thiery, J., Kan, Y. W. & Werner, S. 2008. Impaired liver regeneration in nrf2 knockout mice: Role of ros-mediated insulin/igf-1 resistance. *The EMBO Journal,* 27**,** 212-223.

Biswas, M. & Chan, J. Y. 2012. Gsk3 negatively regulates stress response mediated by transcription factor nrf1, resulting in increased neuronal apoptosis and neurodegeneration. *The FASEB Journal,* 26**,** lb479-lb479.

Biswas, M., Phan, D., Watanabe, M. & Chan, J. Y. 2011. The fbw7 tumor suppressor regulates nuclear factor e2-related factor 1 transcription factor turnover through proteasome-mediated proteolysis. *The Journal of Biological Chemistry,* 286**,** 39282-39289.

Blake, D. J., Singh, A., Kombairaju, P., Malhotra, D., Mariani, T. J., Tuder, R. M., Gabrielson, E. & Biswal, S. 2010. Deletion of keap1 in the lung attenuates acute cigarette smoke–induced oxidative stress and inflammation. *American Journal of Respiratory Cell and Molecular Biology,* 42**,** 524-536.

Blank, V. 2008. Small maf proteins in mammalian gene control: Mere dimerization partners or dynamic transcriptional regulators? *Journal of Molecular Biology,* 376**,** 913-925.

Bratic, A. & Larsson, N.-G. 2013. The role of mitochondria in aging. *The Journal of Clinical Investigation,* 123**,** 951-957.

Buetow, L. & Huang, D. T. 2016. Structural insights into the catalysis and regulation of e3 ubiquitin ligases. *Nature Reviews Molecular Cell Biology,* 17**,** 626.

Bugno, M., Daniel, M., Chepelev, N. L. & Willmore, W. G. 2015. Changing gears in nrf1 research, from mechanisms of regulation to its role in disease and prevention. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms,* 1849**,** 1260-1276.

Campbell, J. D., Alexandrov, A., Kim, J., Wala, J., Berger, A. H., Pedamallu, C. S., Shukla, S. A., Guo, G., Brooks, A. N., Murray, B. A., Imielinski, M., Hu, X., Ling, S., Akbani, R., Rosenberg, M., Cibulskis, C., Ramachandran, A., Collisson, E. A., Kwiatkowski, D. J., Lawrence, M. S., Weinstein, J. N., Verhaak, R. G. W., Wu, C. J., Hammerman, P. S., Cherniack, A. D., Getz, G., Cancer Genome Atlas Research, N., Artyomov, M. N., Schreiber, R., Govindan, R. & Meyerson, M. 2016. Distinct patterns of somatic genome alterations in lung adenocarcinomas and squamous cell carcinomas. *Nature Genetics,* 48**,** 607.

Campbell, L. E. & Proud, C. G. 2002. Differing substrate specificities of members of the dyrk family of arginine-directed protein kinases. *FEBS Letters,* 510**,** 31-36.

Canning, P., J. Sorrell, F. & N. Bullock, A. 2015. Structural basis of keap1 interactions with nrf2. *Free Radical Biology and Medicine,* 88**,** 101-107.

Cerami, E., Gao, J., Dogrusoz, U., Gross, B. E., Sumer, S. O., Aksoy, B. A., Jacobsen, A., Byrne, C. J., Heuer, M. L., Larsson, E., Antipin, Y., Reva, B., Goldberg, A. P., Sander, C. & Schultz, N. 2012. The cbio cancer genomics portal: An open platform for exploring multidimensional cancer genomics data. *Cancer Discovery,* 2**,** 401-404.

Chan, J. Y., Kwong, M., Lu, R., Chang, J., Wang, B., Yen, T. S. & Kan, Y. W. 1998. Targeted disruption of the ubiquitous cnc-bzip transcription factor, nrf-1, results in anemia and embryonic lethality in mice. *The EMBO Journal,* 17**,** 1779-1787.

Chan, K., Han, X.-D. & Kan, Y. W. 2001. An important function of nrf2 in combating oxidative stress: Detoxification of acetaminophen. *Proceedings of the National Academy of Sciences,* 98**,** 4611-4616.

Chan, K., Lu, R., Chang, J. C. & Kan, Y. W. 1996. Nrf2, a member of the nfe2 family of transcription factors, is not essential for murine erythropoiesis, growth, and development. *Proceedings of the National Academy of Sciences,* 93**,** 13943-13948.

Chandel, N. S., S. McClintock, D., Feliciano, C., M. Wood, T., Melendez, J., M. Rodriguez, A. & T. Schumacker, P. 2000. Reactive oxygen species generated at mitochondrial complex

iii stabilize hypoxia-inducible factor-1α during hypoxia a mechanism of o2 sensing. *Journal of Biochemistry* 275**,** 25130-8.

Chang, S., Jiang, X., Zhao, C., Lee, C. & Ferriero, D. M. 2008. E xogenous low dose hydrogen peroxide increases hypoxia-inducible factor-1alpha protein expression and induces preconditioning protection against ischemia in primary cortical neurons. *Neuroscience letters,* 441**,** 134-138.

Chen, J., Solomides, C., Simpkins, F. & Simpkins, H. 2017. The role of nrf2 and atf2 in resistance to platinum-based chemotherapy. *Cancer Chemotherapy and Pharmacology,* 79**,** 369-380.

Chevillard, G. & Blank, V. 2011. Nfe2l3 (nrf3): The cinderella of the cap'n'collar transcription factors. *Cellular and Molecular Life Sciences,* 68**,** 3337-3348.

Chio, I. I. C., Jafarnejad, S. M., Ponz-Sarvise, M., Park, Y., Rivera, K., Palm, W., Wilson, J., Sangar, V., Hao, Y., Öhlund, D., Wright, K., Filippini, D., Lee, E. J., Da Silva, B., Schoepfer, C., Wilkinson, J. E., Buscaglia, J., DeNicola, G. M., Tiriac, H., Hammell, M., Crawford, H. C., Schmidt, E. E., Thompson, C. B., Pappin, D. J., Sonenberg, N. & Tuveson, D. A. 2016. Nrf2 promotes tumor maintenance by modulating mrna translation in pancreatic cancer. *Cell,* 166**,** 963-976.

Cho, K.-J., Seo, J.-M. & Kim, J.-H. 2011. Bioactive lipoxygenase metabolites stimulation of nadph oxidases and reactive oxygen species. *Molecules and Cells,* 32**,** 1-5.

Choi, B.-h. & Kwak, M.-K. 2016. Shadows of nrf2 in cancer: Resistance to chemotherapy. *Current Opinion in Toxicology,* 1**,** 20-28.

Chong, G., Hölscher, C., Liu, Y. & Li, L. 2012. Gsk3: A key target for the development of novel treatments for type 2 diabetes mellitus and alzheimer disease. *Reviews in the neurosciences,* 23**,** 1-11.

Chorley, B. N., Campbell, M. R., Wang, X., Karaca, M., Sambandan, D., Bangura, F., Xue, P., Pi, J., Kleeberger, S. R. & Bell, D. A. 2012. Identification of novel nrf2-regulated genes by chip-seq: Influence on retinoid x receptor alpha. *Nucleic Acids Research,* 40**,** 7416-7429.

Chowdhry, S., Zhang, Y., McMahon, M., Sutherland, C., Cuadrado, A. & Hayes, J. D. 2013. Nrf2 is controlled by two distinct β-trcp recognition motifs in its neh6 domain, one of which can be modulated by gsk-3 activity. *Oncogene,* 32**,** 3765-3781.

Chowdhury, A. M. M. A., Katoh, H., Hatanaka, A., Iwanari, H., Nakamura, N., Hamakubo, T., Natsume, T., Waku, T. & Kobayashi, A. 2017. Multiple regulatory mechanisms of the biological function of nrf3 (nfe2l3) control cancer cell proliferation. *Scientific Reports,* 7**,** 12494.

Chun, S., Johnson, C., Washburn, J., R Cruz-Correa, M., T Dang, D. & Dang, L. 2010. Oncogenic kras modulates mitochondrial metabolism in human colon cancer cells by inducing hif-1α and hif-2α target genes. *Molecular cancer,* 9**,** 293.

Clark, J., Clore, E. L., Zheng, K., Adame, A., Masliah, E. & Simon, D. K. 2010. Oral n-acetyl-cysteine attenuates loss of dopaminergic terminals in α-synuclein overexpressing mice. *PLOS ONE,* 5**,** 12333.

Clarke, J. D., Dashwood, R. H. & Ho, E. 2008. Multi-targeted prevention of cancer by sulforaphane. *Cancer letters,* 269**,** 291-304.

Cleasby, A., Yon, J., Day, P. J., Richardson, C., Tickle, I. J., Williams, P. A., Callahan, J. F., Carr, R., Concha, N., Kerns, J. K., Qi, H., Sweitzer, T., Ward, P. & Davies, T. G. 2014. Structure of the btb domain of keap1 and its interaction with the triterpenoid antagonist cddo. *PLOS ONE,* 9**,** 98896.

Clodfelder-Miller, B., De Sarno, P., Zmijewska, A. A., Song, L. & Jope, R. S. 2005. Physiological and pathological changes in glucose regulate brain akt and glycogen synthase kinase-3. *The Journal of biological chemistry,* 280**,** 39723-39731.

Cloer, E. W., Siesser, P. F., Cousins, E. M., Goldfarb, D., Mowrey, D. D., Harrison, J. S., Weir, S. J., Dokholyan, N. V. & Major, M. B. 2018. P62-dependent phase separation of patient-derived keap1 mutations and nrf2. *Molecular and Cellular Biology*.

Cohen, P. & Frame, S. 2001. The renaissance of gsk3. *Nature Reviews Molecular Cell Biology,* 2**,** 769.

Cohen, P. & Goedert, M. 2004. Gsk3 inhibitors: Development and therapeutic potential. *Nature Reviews Drug Discovery,* 3**,** 479.

Cole, A., Frame, S. & Cohen, P. 2004. Further evidence that the tyrosine phosphorylation of glycogen synthase kinase-3 (gsk3) in mammalian cells is an autophosphorylation event. *Biochemical Journal,* 377**,** 249.

Cole, S. P., Bhardwaj, G., Gerlach, J. H., Mackie, J. E., Grant, C. E., Almquist, K. C., Stewart, A. J., Kurz, E. U., Duncan, A. M. & Deeley, R. G. 1992. Overexpression of a transporter gene in a multidrug-resistant human lung cancer cell line. *Science,* 258**,** 1650-1654.

Cole, S. P. C. 2014. Targeting multidrug resistance protein 1 (mrp1, abcc1): Past, present, and future. *Annual Review of Pharmacology and Toxicology,* 54**,** 95-117.

Cole, S. P. C. & Deeley, R. G. 2006. Transport of glutathione and glutathione conjugates by mrp1. *Trends in Pharmacological Sciences,* 27**,** 438-446.

Collet, J.-F. & Messens, J. 2010. Structure, function, and mechanism of thioredoxin proteins. *Antioxidants & Redox Signaling,* 13**,** 1205-1216.

Connor, K. M., Hempel, N., Nelson, K. K., Dabiri, G., Gamarra, A., Belarmino, J., Van De Water, L., Mian, B. M. & Melendez, J. A. 2007. Manganese superoxide dismutase enhances the invasive and migratory activity of tumor cells. *Cancer Research,* 67**,** 10260-10267.

Cross, D. A. E., Alessi, D. R., Cohen, P., Andjelkovich, M. & Hemmings, B. A. 1995. Inhibition of glycogen synthase kinase-3 by insulin mediated by protein kinase b. *Nature,* 378**,** 785.

Cuadrado, A., Moreno-Murciano, P. & Pedraza-Chaverri, J. 2009. The transcription factor nrf2 as a new therapeutic target in parkinson's disease. *Expert Opinion on Therapeutic Targets,* 13**,** 319-329.

Cui, X. 2012. Reactive oxygen species: The achilles' heel of cancer cells? *Antioxidants & Redox Signaling,* 16**,** 1212-1214.

Cui, X., Li, L., Yan, G., Meng, K., Lin, Z., Nan, Y., Jin, G. & Li, C. 2015. High expression of nqo1 is associated with poor prognosis in serous ovarian carcinoma. *BMC Cancer,* 15**,** 244.

Dajani, R., Fraser, E., Roe, S. M., Young, N., Good, V., Dale, T. C. & Pearl, L. H. 2001. Crystal structure of glycogen synthase kinase 3β: Structural basis for phosphate-primed substrate specificity and autoinhibition. *Cell,* 105**,** 721-732.

Dasari, S. & Tchounwou, P. B. 2014. Cisplatin in cancer therapy: Molecular mechanisms of action. *European journal of pharmacology,* 0**,** 364-378.

Daseul Kim, A., Zhang, R., Xia, H., Kang, K., Jing Piao, M., Hee Maeng, Y., Chang, W. & Won Hyun, J. 2015. Involvement of glutathione and glutathione metabolizing enzymes in human colorectal cancer cell lines and tissues. *Molecular medicine reports,* 12**,** 4314-4319.

de la Parra, C., Walters, B. A., Geter, P. & Schneider, R. J. 2018. Translation initiation factors and their relevance in cancer. *Current Opinion in Genetics & Development,* 48**,** 82-88.

DeNicola, G. M., Karreth, F. A., Humpton, T. J., Gopinathan, A., Wei, C., Frese, K., Mangal, D., Yu, K. H., Yeo, C. J., Calhoun, E. S., Scrimieri, F., Winter, J. M., Hruban, R. H., Iacobuzio-Donahue, C., Kern, S. E., Blair, I. A. & Tuveson, D. A. 2011. Oncogene-induced nrf2 transcription promotes ros detoxification and tumorigenesis. *Nature,* 475**,** 106.

Derjuga, A., Gourley, T. S., Holm, T. M., Heng, H. H. Q., Shivdasani, R. A., Ahmed, R., Andrews, N. C. & Blank, V. 2004. Complexity of cnc transcription factors as revealed by gene targeting of the nrf3 locus. *Molecular and Cellular Biology,* 24**,** 3286-3294.

Dinkova-Kostova, A. T., Kostov, R. V. & Canning, P. 2017. Keap1, the cysteine-based mammalian intracellular sensor for electrophiles and oxidants. *Archives of Biochemistry and Biophysics,* 617**,** 84-93.

Doble, B. W. & Woodgett, J. R. 2003. Gsk-3: Tricks of the trade for a multi-tasking kinase. *Journal of cell science,* 116**,** 1175-1186.

Dorion, S., Lambert, H. & Landry, J. 2002. Activation of the p38 signaling pathway by heat shock involves the dissociation of glutathione s-transferase mu from ask1. *The Journal of biological chemistry,* 277**,** 30792-7.

Dorr, C., Janik, C., Weg, M., Been, R. A., Bader, J., Kang, R., Ng, B., Foran, L., Landman, S. R., O'Sullivan, M. G., Steinbach, M., Sarver, A. L., Silverstein, K. A. T., Largaespada, D. A. & Starr, T. K. 2015. Transposon mutagenesis screen identifies potential lung cancer drivers and cul3 as a tumor suppressor. *Molecular Cancer Research,* 13**,** 1238-1247.

Dostalek, M., Hardy, K., Milne, G., D Morrow, J., Chen, C., J Gonzalez, F., Gu, J., Ding, X., A Johnson, D., A Johnson, J., V Martin, M. & Guengerich, F. 2008. Development of oxidative stress by cytochrome p450 induction in rodents is selective for barbiturates and related to loss of pyridine nucleotide-dependent protective systems. *Journal of Biological Chemistry,* 283**,** 17147-57.

Dunnill, C., Patton, T., Brennan, J., Barrett, J., Dryden, M., Cooke, J., Leaper, D. & Georgopoulos, N. T. 2017. Reactive oxygen species (ros) and wound healing: The functional role of ros and emerging ros-modulating technologies for augmentation of the healing process. *International Wound Journal,* 14**,** 89-96.

El-Ashmawy, M., Delgado, O., Cardentey, A., Wright, W. E. & Shay, J. W. 2014. Cddo-me protects normal lung and breast epithelial cells but not cancer cells from radiation. *PLOS ONE,* 9**,** 115600.

El-Osta, H. & Circu, M. L. 2016. Mitochondrial ros and apoptosis. *In:* BUHLMAN, L. M. (ed.) *Mitochondrial mechanisms of degeneration and repair in parkinson's disease.* Cham: Springer International Publishing.

Elgebaly, S. A., Hall, I. H., Lee, K. H., Sumida, Y., Imakura, Y. & Wu, R. Y. 1979. Antitumor agents xxxv: Effects of brusatol, bruceoside a, and bruceantin on p-388 lymphocytic leukemia cell respiration. *Journal of Pharmaceutical Sciences,* 68**,** 887-890.

Embi, N., Rylatt, D. B. & Cohen, P. 1980. Glycogen synthase kinase-3 from rabbit skeletal muscle. *European Journal of Biochemistry,* 107**,** 519-527.

Emir, U. E., Raatz, S., McPherson, S., Hodges, J. S., Torkelson, C., Tawfik, P., White, T. & Terpstra, M. 2011. Noninvasive quantification of ascorbate and glutathione concentration in the elderly human brain. *NMR in biomedicine,* 24**,** 888-894.

Enomoto, A., Itoh, K., Nagayoshi, E., Haruta, J., Kimura, T., O'Connor, T., Harada, T. & Yamamoto, M. 2001. High sensitivity of nrf2 knockout mice to acetaminophen hepatotoxicity associated with decreased expression of are-regulated drug metabolizing enzymes and antioxidant genes. *Toxicological Sciences,* 59**,** 169-177.

Erickson, A., M, Nevarea, Z., J Gipp, J. & Timothy Mulcahy, R. 2002. Identification of a variant antioxidant response element in the promoter of the human glutamate-cysteine ligase modifier subunit gene. Revision of the are consensus sequence. *The Journal of biological chemistry,* 277**,** 30730-7.

Erkens, R., Kramer, C. M., Lückstädt, W., Panknin, C., Krause, L., Weidenbach, M., Dirzka, J., Krenz, T., Mergia, E., Suvorava, T., Kelm, M. & Cortese-Krott, M. M. 2015. Left ventricular diastolic dysfunction in nrf2 knock out mice is associated with cardiac hypertrophy, decreased expression of serca2a, and preserved endothelial function. *Free Radical Biology and Medicine,* 89**,** 906-917.

Evans, M. 2013. Lung cancer: Needs assessment, treatment and therapies. *British journal of nursing (Mark Allen Publishing),* 22**,** S15-22.

Fahey, J. W., Dinkova-Kostova, A. T., Stephenson, K. K. & Talalay, P. 2004. The “prochaska” microtiter plate bioassay for inducers of nqo1. *Methods in Enzymology,* 382**,** 243-258.

Fan, Z., Wirth, A. K., Chen, D., Wruck, C. J., Rauh, M., Buchfelder, M. & Savaskan, N. 2017. Nrf2-keap1 pathway promotes cell proliferation and diminishes ferroptosis. *Oncogenesis,* 6**,** 371.

Faraonio, R., Vergara, P., Di Marzo, D., Giovanna Pierantoni, M., Napolitano, M., Russo, T. & Cimino, F. 2007. P53 suppresses the nrf2-dependent transcription of antioxidant response genes. *The Journal of biological chemistry,* 281**,** 39776-84.

Favreau, L. & Pickett, C. 1993. Transcriptional regulation of the rat nad(p)h:Quinone reductase gene. Characterization of a DNA-protein interaction at the antioxidant responsive element and induction by 12-o-tetradecanoylphorbol 13-acetate. *The Journal of biological chemistry,* 268**,** 19875-81.

Feitelson, M. A., Arzumanyan, A., Kulathinal, R. J., Blain, S. W., Holcombe, R. F., Mahajna, J., Marino, M., Martinez-Chantar, M. L., Nawroth, R., Sanchez-Garcia, I., Sharma, D., Saxena, N. K., Singh, N., Vlachostergios, P. J., Guo, S., Honoki, K., Fujii, H., Georgakilas, A. G., Bilsland, A., Amedei, A., Niccolai, E., Amin, A., Ashraf, S. S., Boosani, C. S., Guha, G., Ciriolo, M. R., Aquilano, K., Chen, S., Mohammed, S. I., Azmi, A. S., Bhakta, D., Halicka, D., Keith, W. N. & Nowsheen, S. 2015. Sustained proliferation in cancer: Mechanisms and novel therapeutic targets. *Seminars in Cancer Biology,* 35**,** S25-S54.

Finkel T. Reactive Oxygen Species and Signal Transduction.2001. *IUBMB Life*, 52(1), 3-6.

Finkel T. Signal transduction by reactive oxygen species. 2011. *The Journal of cell biology*, 194(1), 7-15.

Foerster, E. C., Fährenkemper, T., Rabe, U., Graf, P. & Sies, H. 1981. Peroxisomal fatty acid oxidation as detected by h₂o₂ production in intact perfused rat liver. *Biochemical Journal,* 196**,** 705-712.

Fourquet, S., Guerois, R., Biard, D. & Toledano, M. B. 2010. Activation of nrf2 by nitrosative agents and h(2)o(2) involves keap1 disulfide formation. *The Journal of Biological Chemistry,* 285**,** 8463-8471.

Freland, L. & Beaulieu, J.-M. 2012. Inhibition of gsk3 by lithium, from single molecules to signaling networks. *Frontiers in Molecular Neuroscience,* 5**,** 14.

Fukutomi, T., Takagi, K., Mizushima, T., Ohuchi, N. & Yamamoto, M. 2014. Kinetic, thermodynamic, and structural characterizations of the association between nrf2-dlgex degron and keap1. *Molecular and Cellular Biology,* 34**,** 832-846.

Futreal, P. A., Coin, L., Marshall, M., Down, T., Hubbard, T., Wooster, R., Rahman, N. & Stratton, M. R. 2004. A census of human cancer genes. *Nature Reviews Cancer,* 4**,** 177.

Gabriel, D., Roedl, D., Gordon, L. B. & Djabali, K. 2015. Sulforaphane enhances progerin clearance in hutchinson–gilford progeria fibroblasts. *Aging Cell,* 14**,** 78-91.

Galanis, A., Pappa, A., Giannakakis, A., Lanitis, E., Dangaj, D. & Sandaltzopoulos, R. 2008. Reactive oxygen species and hif-1 signalling in cancer. *Cancer Letters,* 266**,** 12-20.

Gamcsik, M. P., Kasibhatla, M. S., Teeter, S. D. & Colvin, O. M. 2012. Glutathione levels in human tumors. *Biomarkers,* 17**,** 671-691.

Gao, C., Chen, G., Romero, G., Moschos, S., Xu, X. & Hu, J. 2014. Induction of gsk3β-β-trcp interaction is required for late phase stabilization of β-catenin in canonical wnt signaling. *The Journal of Biological Chemistry,* 289**,** 7099-7108.

Gao, J., Aksoy, B. A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S. O., Sun, Y., Jacobsen, A., Sinha, R., Larsson, E., Cerami, E., Sander, C. & Schultz, N. 2013. Integrative analysis of complex cancer genomics and clinical profiles using the cbioportal. *Science Signaling,* 6**,** pl1-pl1.

Gaziano, J. M., Glynn, R. J., Christen, W. G., Kurth, T., Belanger, C., MacFadyen, J., Bubes, V., Manson, J. E., Sesso, H. D. & Buring, J. E. 2009. Vitamins e and c in the prevention of prostate and total cancer in men: The physicians' health study ii, a randomized controlled trial. *JAMA : the journal of the American Medical Association,* 301**,** 52-62.

Giacco, F. & Brownlee, M. 2010. Oxidative stress and diabetic complications. *Circulation research,* 107**,** 1058-1070.

Gibbons, D. L., Byers, L. A. & Kurie, J. M. 2014. Smoking, p53 mutation, and lung cancer. *Molecular Cancer Research,* 12**,** 3-13.

Giese Karl, P. 2009. Gsk-3: A key player in neurodegeneration and memory. *IUBMB Life,* 61**,** 516-521.

Go Y-M, Chandler JD, Jones DP. 2015. The cysteine proteome. *Free radical biology & medicine,* 84, 227-45.

Gomez-Sintes R, Hernandez F, Lucas J, Avila J. GSK-3 Mouse Models to Study Neuronal Apoptosis and Neurodegeneration. 2011. *Frontiers in Molecular Neuroscience*, 4(45).

Gorelik, M., Orlicky, S., Sartori, M. A., Tang, X., Marcon, E., Kurinov, I., Greenblatt, J. F., Tyers, M., Moffat, J., Sicheri, F. & Sidhu, S. S. 2016. Inhibition of scf ubiquitin ligases by engineered ubiquitin variants that target the cul1 binding site on the skp1–f-box interface. *Proceedings of the National Academy of Sciences,* 113**,** 3527-3532.

Gottesman, M. M. 2002. Mechanisms of cancer drug resistance. *Annual Review of Medicine,* 53**,** 615-627.

Guo, Y., Yu, S., Zhang, C. & Kong, A.-N. T. 2015. Epigenetic regulation of keap1-nrf2 signaling. *Free radical biology & medicine,* 88**,** 337-349.

Habib, E., Linher-Melville, K., Lin, H.-X. & Singh, G. 2015. Expression of xct and activity of system xc− are regulated by nrf2 in human breast cancer cells in response to oxidative stress. *Redox Biology,* 5**,** 33-42.

Hall, I. H., Lee, K. H., Elgebaly, S. A., Imakura, Y., Sumida, Y. & Wu, R. Y. 1979. Antitumor agents xxxiv: Mechanism of action of bruceoside a and brusatol on nucleic acid metabolism of p-388 lymphocytic leukemia cells. *Journal of Pharmaceutical Sciences,* 68**,** 883-887.

Halliwell, B. 2006. Phagocyte-derived reactive species: Salvation or suicide? *Trends in Biochemical Sciences,* 31**,** 509-15.

Halliwell, B. 2007. Biochemistry of oxidative stress. *Biochemical Society Transactions,* 35**,** 1147-1150.

Hanada, N., Takahata, T., Zhou, Q., Ye, X., Sun, R., Itoh, J., Ishiguro, A., Kijima, H., Mimura, J., Itoh, K., Fukuda, S. & Saijo, Y. 2012. Methylation of the keap1 gene promoter region in human colorectal cancer. *BMC Cancer,* 12**,** 66-66.

Hanahan, D. & Weinberg, R. A. 2000. The hallmarks of cancer. *Cell,* 100**,** 57-70.

Hanahan, D. & Weinberg, Robert A. 2011. Hallmarks of cancer: The next generation. *Cell,* 144**,** 646-674.

Hansen, R. E., Roth, D. & Winther, J. R. 2009. Quantifying the global cellular thiol–disulfide status. *Proceedings of the National Academy of Sciences of the United States of America,* 106**,** 422-427.

Harman, D. 1956. Aging: A theory based on free radical and radiation chemistry. *Journal of Gerontology,* 11**,** 298-300.

Hast, B. E., Cloer, E. W., Goldfarb, D., Li, H., Siesser, P. F., Yan, F., Walter, V., Zheng, N., Hayes, D. N. & Major, M. B. 2014. Cancer-derived mutations in keap1 impair nrf2 degradation but not ubiquitination. *Cancer Research,* 74**,** 808-817.

Hast, B. E., Goldfarb, D., Mulvaney, K. M., Hast, M. A., Siesser, P. F., Yan, F., Hayes, D. N. & Major, M. B. 2013. Proteomic analysis of ubiquitin ligase keap1 reveals associated proteins that inhibit nrf2 ubiquitination. *Cancer Research,* 73**,** 2199-2210.

Hastie, C. J., McLauchlan, H. J. & Cohen, P. 2006. Assay of protein kinases using radiolabeled atp: A protocol. *Nature Protocols,* 1**,** 968.

Hayes, J. D., Chanas, S. A., Henderson, C. J., McMahon, M., Sun, C., Moffat, G. J., Wolf, C. R. & Yamamoto, M. 2000. The nrf2 transcription factor contributes both to the basal expression of glutathione s-transferases in mouse liver and to their induction by the chemopreventive synthetic antioxidants, butylated hydroxyanisole and ethoxyquin. *Biochemical Society Transactions,* 28**,** 33-41.

Hayes, J. D. & Dinkova-Kostova, A. T. 2014. The nrf2 regulatory network provides an interface between redox and intermediary metabolism. *Trends in Biochemical Sciences,* 39**,** 199-218.

Hayes, J. D., Dinkova-Kostova, A. T. & McMahon, M. 2009. Cross-talk between transcription factors ahr and nrf2: Lessons for cancer chemoprevention from dioxin. *Toxicological Sciences,* 111**,** 199-201.

Hayes, J. D., Flanagan, J. U. & Jowsey, I. R. 2005. Glutathione transferases. *Annual Review of Pharmacology and Toxicology,* 45**,** 51-88.

Hayes, J. D. & McMahon, M. 2009. Nrf2 and keap1 mutations: Permanent activation of an adaptive response in cancer. *Trends in Biochemical Sciences,* 34**,** 176-188.

Hensing, T., MD 2014. A personalized treatment for lung cancer: Molecular pathways, targeted therapies and genomic characterization. *In:* MALTSEV, N., RZHETSKII, A. I. & GILLIAM, T. C. (eds.) *Systems analysis of human multigene disorders.*

Hernández, F., Barreda, E. G. d., Fuster-Matanzo, A., Goñi-Oliver, P., Lucas, J. J. & Avila, J. 2009. The role of gsk3 in alzheimer disease. *Brain Research Bulletin,* 80**,** 248-250.

Herzig, S. & Shaw, R. J. 2018. Ampk: Guardian of metabolism and mitochondrial homeostasis. *Nature reviews. Molecular cell biology,* 19**,** 121-135.

Hoeflich, K. P., Luo, J., Rubie, E. A., Tsao, M.-S., Jin, O. & Woodgett, J. R. 2000. Requirement for glycogen synthase kinase-3β in cell survival and nf-κb activation. *Nature,* 406**,** 86.

Homma, S., Ishii, Y., Morishima, Y., Yamadori, T., Matsuno, Y., Haraguchi, N., Kikuchi, N., Satoh, H., Sakamoto, T., Hizawa, N., Itoh, K. & Yamamoto, M. 2009. Nrf2 enhances cell proliferation and resistance to anticancer drugs in human lung cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research,* 15**,** 3423-32.

Hong, D. S., Kurzrock, R., Supko, J. G., He, X., Naing, A., Wheler, J., Lawrence, D., Eder, J. P., Meyer, C. J., Ferguson, D. A., Mier, J., Konopleva, M., Konoplev, S., Andreeff, M., Kufe, D., Lazarus, H., Shapiro, G. I. & Dezube, B. J. 2012. A phase i first-in-human trial of bardoxolone methyl in patients with advanced solid tumors and lymphomas. *Clinical cancer research : an official journal of the American Association for Cancer Research,* 18**,** 3396-3406.

Horita, N., Nagashima, A., Nakashima, K., Shibata, Y., Ito, K., Goto, A., Yamanaka, T. & Kaneko, T. 2017. The best platinum regimens for chemo-naive incurable non-small cell lung cancer: Network meta-analysis. *Scientific Reports,* 7**,** 13185.

Hourihan, J. M., Kenna, J. G. & Hayes, J. D. 2012. The gasotransmitter hydrogen sulfide induces nrf2-target genes by inactivating the keap1 ubiquitin ligase substrate adaptor through formation of a disulfide bond between cys-226 and cys-613. *Antioxidants & Redox Signaling,* 19**,** 465-481.

Hrycay, E. G. & Bandiera, S. M. 2015. Chapter two - involvement of cytochrome p450 in reactive oxygen species formation and cancer. *In:* HARDWICK, J. P. (ed.) *Advances in pharmacology.* Academic Press.

Hu, C., Eggler, A. L., Mesecar, A. D. & van Breemen, R. B. 2011a. Modification of keap1 cysteine residues by sulforaphane. *Chemical research in toxicology,* 24**,** 515-521.

Hu, R., Xu, C., Shen, G., Jain, M. R., Khor, T. O., Gopalkrishnan, A., Lin, W., Reddy, B., Chan, J. Y. & Kong, A.-N. T. 2006. Identification of nrf2-regulated genes induced by chemopreventive isothiocyanate peitc by oligonucleotide microarray. *Life Sciences,* 79**,** 1944-1955.

Hu, Y., Lu, W., Chen, G., Wang, P., Chen, Z., Zhou, Y., Ogasawara, M., Trachootham, D., Feng, L., Pelicano, H., J Chiao, P., J Keating, M., Garcia-Manero, G. & Huang, P. 2011b. K-ras g12v transformation leads to mitochondrial dysfunction and a metabolic switch from oxidative phosphorylation to glycolysis. *Cell research,* 22**,** 399-412.

Huang, J., Xiao, Y., Xu, A. & Zhou, Z. 2016. Neutrophils in type 1 diabetes. *Journal of Diabetes Investigation,* 7**,** 652-663.

Hughes, K., Nikolakaki, E., Plyte, S. E., Totty, N. F. & Woodgett, J. R. 1993. Modulation of the glycogen synthase kinase-3 family by tyrosine phosphorylation. *The EMBO journal,* 12**,** 803-808.

Huppke, P., Weissbach, S., Church, J. A., Schnur, R., Krusen, M., Dreha-Kulaczewski, S., Kühn-Velten, W. N., Wolf, A., Huppke, B., Millan, F., Begtrup, A., Almusafri, F., Thiele, H., Altmüller, J., Nürnberg, P., Müller, M. & Gärtner, J. 2017. Activating de novo mutations in nfe2l2 encoding nrf2 cause a multisystem disorder. *Nature Communications,* 8**,** 818.

Hurtado DE, Molina-Porcel L, Carroll JC, Macdonald C, Aboagye AK, Trojanowski JQ, et al. 2012. Selectively silencing GSK-3 isoforms reduces plaques and tangles in mouse models of Alzheimer's disease. T*he Journal of neuroscience : the official journal of the Society for Neuroscience*;32(21),7392-402.

Ichimura, Y., Waguri, S., Sou, Y.-S., Kageyama, S., Hasegawa, J., Ishimura, R., Saito, T., Yang, Y., Kouno, T., Fukutomi, T., Hoshii, T., Hirao, A., Takagi, K., Mizushima, T., Motohashi, H., Lee, M.-S., Yoshimori, T., Tanaka, K., Yamamoto, M. & Komatsu, M. 2013. Phosphorylation of p62 activates the keap1-nrf2 pathway during selective autophagy. *Molecular cell,* 51**,** 618-631.

Ighodaro, O. M. & Akinloye, O. A. 2017. First line defence antioxidants-superoxide dismutase (sod), catalase (cat) and glutathione peroxidase (gpx): Their fundamental role in the entire antioxidant defence grid. *Alexandria Journal of Medicine*.

Innamorato, N. G., Rojo, A. I., García-Yagüe, Á. J., Yamamoto, M., de Ceballos, M. L. & Cuadrado, A. 2008. The transcription factor nrf2 is a therapeutic target against brain inflammation. *The Journal of Immunology,* 181**,** 680-689.

Itoh, K., Chiba, T., Takahashi, S., Ishii, T., Igarashi, K., Katoh, Y., Oyake, T., Hayashi, N., Satoh, K., Hatayama, I., Yamamoto, M. & Nabeshima, Y.-i. 1997. An nrf2/small maf heterodimer mediates the induction of phase ii detoxifying enzyme genes through antioxidant response elements. *Biochemical and Biophysical Research Communications,* 236**,** 313-322.

Itoh, K., Igarashi, K., Hayashi, N., Nishizawa, M. & Yamamoto, M. 1995. Cloning and characterization of a novel erythroid cell-derived cnc family transcription factor heterodimerizing with the small maf family proteins. *Molecular and Cellular Biology,* 15**,** 4184-4193.

Itoh, K., Wakabayashi, N., Katoh, Y., Ishii, T., Igarashi, K., Engel, J. D. & Yamamoto, M. 1999. Keap1 represses nuclear activation of antioxidant responsive elements by nrf2 through binding to the amino-terminal neh2 domain. *Genes & Development,* 13**,** 76-86.

Itoh, K., Wakabayashi, N., Katoh, Y., Ishii, T., O'Connor, T. & Yamamoto, M. 2003. Keap1 regulates both cytoplasmic-nuclear shuttling and degradation of nrf2 in response to electrophiles. *Genes to Cells,* 8**,** 379-391.

Jaiswal, A. K. 2000. Regulation of genes encoding nad(p)h:Quinone oxidoreductases. *Free Radical Biology and Medicine,* 29**,** 254-262.

Jarvi MT, Patterson MS, Wilson BC. 2012. Insights into photodynamic therapy dosimetry: simultaneous singlet oxygen luminescence and photosensitizer photobleaching measurements*. Biophysical journal*, 102(3), 661-71.

Ji, L., Li, H., Gao, P., Shang, G., Zhang, D. D., Zhang, N. & Jiang, T. 2013. Nrf2 pathway regulates multidrug-resistance-associated protein 1 in small cell lung cancer. *PLOS ONE,* 8**,** 63404.

Jiang, L., Thor, D. S. & Jeffrey, A. J. 2004. Genetic dissection of systemic autoimmune disease in nrf2-deficient mice. *Physiological Genomics,* 18**,** 261-272.

Joo, M. S., Kim, W. D., Lee, K. Y., Kim, J. H., Koo, J. H. & Kim, S. G. 2016. Ampk facilitates nuclear accumulation of nrf2 by phosphorylating at serine 550. *Molecular and Cellular Biology,* 36**,** 1931-1942.

Jope, R. S. & Johnson, G. V. W. 2004. The glamour and gloom of glycogen synthase kinase-3. *Trends in Biochemical Sciences,* 29**,** 95-102.

Jope, R. S., Yuskaitis, C. J. & Beurel, E. 2007. Glycogen synthase kinase-3 (gsk3): Inflammation, diseases, and therapeutics. *Neurochemical research,* 32**,** 577-595.

Kang, M.-I., Kobayashi, A., Wakabayashi, N., Kim, S.-G. & Yamamoto, M. 2004. Scaffolding of keap1 to the actin cytoskeleton controls the function of nrf2 as key regulator of cytoprotective phase 2 genes. *Proceedings of the National Academy of Sciences of the United States of America,* 101**,** 2046-2051.

Kaplowitz, N., Aw, T. & Ookhtens, M. 1985. The regulation of hepatic gsh. *Annual review of pharmacology and toxicology,* 25**,** 715-44.

Katoh, Y., Itoh, K., Yoshida, E., Miyagishi, M., Fukamizu, A. & Yamamoto, M. 2002. Two domains of nrf2 cooperatively bind cbp, a creb binding protein, and synergistically activate transcription. *Genes to Cells,* 6**,** 857-868.

Katsuoka, F., Motohashi, H., Ishii, T., Aburatani, H., Engel, J. D. & Yamamoto, M. 2005. Genetic evidence that small maf proteins are essential for the activation of antioxidant response element-dependent genes. *Molecular and Cellular Biology,* 25**,** 8044-8051.

Kawai, Y., Garduño, L., Malloy, M., Yang, J. & J Arinze, I. 2010. Acetylation-deacetylation of the transcription factor nrf2 (nuclear factor erythroid 2-related factor 2) regulates its transcriptional activity and nucleocytoplasmic localization. *The Journal of biological chemistry,* 286**,** 7629-40.

Kensler, T. W., Egner, P. A., Agyeman, A. S., Visvanathan, K., Groopman, J. D., Chen, J.-G., Chen, T.-Y., Fahey, J. W. & Talalay, P. 2013. Keap1-nrf2 signaling: A target for cancer prevention by sulforaphane. *Topics in current chemistry,* 329**,** 163-177.

Khor, T. O., Huang, M.-T., Prawan, A., Liu, Y., Hao, X., Yu, S., Cheung, W. K. L., Chan, J. Y., Reddy, B. S., Yang, C. S. & Kong, A.-N. 2008. Increased susceptibility of nrf2 knockout mice to colitis-associated colorectal cancer. *Cancer prevention research (Philadelphia, Pa.),* 1**,** 187-191.

Kim, E.-H., Deng, C., Sporn, M. B., Royce, D. B., Risingsong, R., Williams, C. R. & Liby, K. T. 2012. Cddo-methyl ester delays breast cancer development in brca1-mutated mice. *Cancer prevention research (Philadelphia, Pa.),* 5**,** 89-97.

Kim, Y.-W. & Byzova, T. V. 2014. Oxidative stress in angiogenesis and vascular disease. *Blood,* 123**,** 625-631.

Kobayashi, A., Ito, E., Toki, T., Kogame, K., Takahashi, S., Igarashi, K., Hayashi, N. & Yamamoto, M. 1999. Molecular cloning and functional characterization of a new cap'n' collar family transcription factor nrf3. *Journal of Biological Chemistry,* 274**,** 6443-6452.

Kobayashi, M., Itoh, K., Suzuki, T., Osanai, H., Nishikawa, K., Katoh, Y., Takagi, Y. & Yamamoto, M. 2002. Identification of the interactive interface and phylogenic conservation of the nrf2-keap1 system. *Genes to Cells,* 7**,** 807-820.

Kobayashi, M., Li, L., Iwamoto, N., Nakajima-Takagi, Y., Kaneko, H., Nakayama, Y., Eguchi, M., Wada, Y., Kumagai, Y. & Yamamoto, M. 2009. The antioxidant defense system keap1-nrf2 comprises a multiple sensing mechanism for responding to a wide range of chemical compounds. *Molecular and Cellular Biology,* 29**,** 493-502.

Komatsu, M., Kurokawa, H., Waguri, S., Taguchi, K., Kobayashi, A., Ichimura, Y., Sou, Y.-S., Ueno, I., Sakamoto, A., Tong, K. I., Kim, M., Nishito, Y., Iemura, S.-i., Natsume, T., Ueno, T., Kominami, E., Motohashi, H., Tanaka, K. & Yamamoto, M. 2010. The selective autophagy substrate p62 activates the stress responsive transcription factor nrf2 through inactivation of keap1. *Nature Cell Biology,* 12**,** 213.

Kondratiuk I, Devijver H, Lechat B, Van Leuven F, Kaczmarek L, Filipkowski RK .2013. Glycogen synthase kinase-3beta affects size of dentate gyrus and species-typical behavioral tasks in transgenic and knockout mice. *Behavioural Brain Research*, 248,46-50.

Kong, B., Qia, C., Erkan, M., Kleeff, J. & Michalski, C. 2013. Overview on how oncogenic kras promotes pancreatic carcinogenesis by inducing low intracellular ros levels. *Frontiers in physiology,* 4**,** 246.

Koppula, P., Zhang, Y., Shi, J., Li, W. & Gan, B. 2017. The glutamate/cystine antiporter slc7a11/xct enhances cancer cell dependency on glucose by exporting glutamate. *Journal of Biological Chemistry,* 292**,** 14240-14249.

Koppula, P., Zhang, Y., Zhuang, L. & Gan, B. 2018. Amino acid transporter slc7a11/xct at the crossroads of regulating redox homeostasis and nutrient dependency of cancer. *Cancer Communications,* 38**,** 12.

Kotkow, K. J. & Orkin, S. H. 1995. Dependence of globin gene expression in mouse erythroleukemia cells on the nf-e2 heterodimer. *Molecular and Cellular Biology,* 15**,** 4640-4647.

Kramarenko GG, Hummel SG, Martin SM, Buettner GR. Ascorbate reacts with singlet oxygen to produce hydrogen peroxide.2006. *Photochemistry and photobiology*, 82(6),1634-7.

Kremer, A., Louis, J., Jaworski, T. & Vanleuven, F. 2011. Gsk3 and alzheimer's disease: Facts and fiction…. *Frontiers in Molecular Neuroscience,* 4**,** 17.

Kubben, N., Zhang, W., Wang, L., Voss, T. C., Yang, J., Qu, J., Liu, G.-H. & Misteli, T. 2016. Repression of the antioxidant nrf2 pathway in premature aging. *Cell,* 165**,** 1361-1374.

Kudin, A., Debska-Vielhaber, G. & Kunz, W. 2005. Characterization of superoxide production sites in isolated rat brain and skeletal muscle mitochondria. *Biomedicine and Pharmacotherapy,* 59**,** 163-8.

Kuosmanen SM, Sihvola V, Kansanen E, Kaikkonen MU, Levonen A-L. 2018. MicroRNAs mediate the senescence-associated decline of NRF2 in endothelial cells. *Redox biology*,18,77-83.

Kwak, M.-K., Itoh, K., Yamamoto, M. & Kensler, T. W. 2002. Enhanced expression of the transcription factor nrf2 by cancer chemopreventive agents: Role of antioxidant response element-like sequences in the nrf2 promoter. *Molecular and Cellular Biology,* 22**,** 2883-2892.

Lan, D., Wang, L., He, R., Ma, J., Bin, Y., Chi, X., Chen, G. & Cai, Z. 2018. Exogenous glutathione contributes to cisplatin resistance in lung cancer a549 cells. *American Journal of Translational Research,* 10**,** 1295-1309.

Laplante, M. & Sabatini, D. M. 2012. Mtor signaling in growth control and disease. *Cell,* 149**,** 274-293.

Lassot, I., Ségéral, E., Berlioz-Torrent, C., Durand, H., Groussin, L., Hai, T., Benarous, R. & Margottin-Goguet, F. 2001. Atf4 degradation relies on a phosphorylation-dependent interaction with the scf(betatrcp) ubiquitin ligase. *Molecular and cellular biology,* 21**,** 2192-2202.

Lastres-Becker, I., García-Yagüe, A. J., Scannevin, R. H., Casarejos, M. J., Kügler, S., Rábano, A. & Cuadrado, A. 2016. Repurposing the nrf2 activator dimethyl fumarate as therapy against synucleinopathy in parkinson's disease. *Antioxidants & Redox Signaling,* 25**,** 61-77.

Leroy, B., Girard, L., Hollestelle, A., Minna, J. D., Gazdar, A. F. & Soussi, T. 2014. Analysis of tp53 mutation status in human cancer cell lines: A reassessment. *Human mutation,* 35**,** 756-765.

Li, N., Ragheb, K., Lawler, G., Sturgis, J., Rajwa, B., Melendez, J. & Robinson, J. 2003. Mitochondrial complex i inhibitor rotenone induces apoptosis through enhancing mitochondrial reactive oxygen species production. *Journal of Biological Chemistry,* 278**,** 8516-25.

Li, X., Zhang, D., Hannink, M. & Beamer, L. 2005. Crystal structure of the kelch domain of human keap1. *Journal of Biological Chemistry,* 279**,** 54750-8.

Liochev, S. I. 2013. Reactive oxygen species and the free radical theory of aging. *Free Radical Biology and Medicine,* 60**,** 1-4.

Liou, G.-Y. & Storz, P. 2010. Reactive oxygen species in cancer. *Free Radical Research,* 44**,** 479-96.

Lisek, K., Campaner, E., Ciani, Y., Walerych, D. & Del Sal, G. 2018. Mutant p53 tunes the nrf2-dependent antioxidant response to support survival of cancer cells. *Oncotarget,* 9**,** 20508-20523.

Lister, A., Nedjadi, T., Kitteringham, N. R., Campbell, F., Costello, E., Lloyd, B., Copple, I. M., Williams, S., Owen, A., Neoptolemos, J. P., Goldring, C. E. & Park, B. K. 2011. Nrf2 is overexpressed in pancreatic cancer: Implications for cell proliferation and therapy. *Molecular Cancer,* 10**,** 37-37.

Liu, C., Li, Y., Semenov, M., Han, C., Baeg, G.-H., Tan, Y., Zhang, Z., Lin, X. & He, X. 2002. Control of β-catenin phosphorylation/degradation by a dual-kinase mechanism. *Cell,* 108**,** 837-847.

Loignon, M., Miao, W., Hu, L., Bier, A., Bismar, T. A., Scrivens, P. J., Mann, K., Basik, M., Bouchard, A., Fiset, P. O., Batist, Z. & Batist, G. 2009. Cul3 overexpression depletes nrf2 in breast cancer and is associated with sensitivity to carcinogens, to oxidative stress, and to chemotherapy. *Molecular Cancer Therapeutics,* 8**,** 2432-2440.

Lu, J. & Holmgren, A. 2014. The thioredoxin antioxidant system. *Free Radical Biology and Medicine,* 66**,** 75-87.

Lu, K., Alcivar, A. L., Ma, J., Foo, T. K., Zywea, S., Huo, Y., Kensler, T. W., Gatza, M. L. & Xia, B. 2016. 307 - dpp3 in nrf2 signaling and breast cancer. *Free Radical Biology and Medicine,* 100**,** S132.

Lu, K., Alcivar, A. L., Ma, J., Foo, T. K., Zywea, S., Mahdi, A., Huo, Y., Kensler, T. W., Gatza, M. L. & Xia, B. 2017. Nrf2 induction supporting breast cancer cell survival is enabled by oxidative stress-induced dpp3-keap1 interaction. *Cancer research,* 77**,** 2881-2892.

Lu, S. C. 2009. Regulation of glutathione synthesis. *Molecular aspects of medicine,* 30**,** 42-59.
Lu, S. C. 2013. Glutathione synthesis. *Biochimica et biophysica acta,* 1830**,** 3143-3153.
Luo, S. Y. & Lam, D. C. 2013. Oncogenic driver mutations in lung cancer. *Translational Respiratory Medicine,* 1**,** 6.
Lv, E., Deng, J., Yu, Y., Wang, Y., Gong, X., Jia, J. & Wang, X. 2015. Nrf2-are signals mediated the anti-oxidative action of electroacupuncture in an mptp mouse model of parkinson's disease. *Free Radical Research,* 49**,** 1296-1307.
Ma, Q., Kinneer, K., Bi, Y., Y Chan, J. & Wai Kan, Y. 2004. Induction of murine nad(p)h:Quinone oxidoreductase by 2,3,7,8-tetrachlorodibenzo-p-dioxin requires the cnc (cap 'n' collar) basic leucine zipper transcription factor nrf2 (nuclear factor erythroid 2-related factor 2): Cross-interaction between ahr (aryl hydrocarbon receptor) and nrf2 signal transduction. *The Biochemical journal,* 377**,** 205-13.
MacAulay, K., Doble, B. W., Patel, S., Hansotia, T., Sinclair, E. M., Drucker, D. J., Nagy, A. & Woodgett, J. R. 2007. Glycogen synthase kinase 3α-specific regulation of murine hepatic glycogen metabolism. *Cell Metabolism,* 6**,** 329-337.
MacLeod, A. K., McMahon, M., Plummer, S. M., Higgins, L. G., Penning, T. M., Igarashi, K. & Hayes, J. D. 2009. Characterization of the cancer chemopreventive nrf2-dependent gene battery in human keratinocytes: Demonstration that the keap1–nrf2 pathway, and not the bach1–nrf2 pathway, controls cytoprotection against electrophiles as well as redox-cycling compounds. *Carcinogenesis,* 30**,** 1571-1580.
Mahaffey, C. M., Mahaffey, N. C., Holland, W., Zhang, H., Gandara, D. R., Mack, P. C. & Forman, H. J. 2012. Aberrant regulation of the mrp3 gene in non-small cell lung carcinoma. *Journal of Thoracic Oncology,* 7**,** 34-39.
Malloy, M. T., McIntosh, D. J., Walters, T. S., Flores, A., Goodwin, J. S. & Arinze, I. J. 2013. Trafficking of the transcription factor nrf2 to promyelocytic leukemia-nuclear bodies: Implications for degradation of nrf2 in the nucleus. *The Journal of Biological Chemistry,* 288**,** 14569-14583.
Margis, R., Dunand, C., K Teixeira, F. & Margis-Pinheiro, M. 2008. Glutathione peroxidase family–an evolutionary overview. *The FEBS journal,* 275**,** 3959-70.
Marin, O., Bustos, V. H., Cesaro, L., Meggio, F., Pagano, M. A., Antonelli, M., Allende, C. C., Pinna, L. A. & Allende, J. E. 2003. A noncanonical sequence phosphorylated by casein kinase 1 in β-catenin may play a role in casein kinase 1 targeting of important signaling proteins. *Proceedings of the National Academy of Sciences,* 100**,** 10193-10200.
Martin, P., Leighl, N. B., Tsao, M.-S. & Shepherd, F. A. 2013. Kras mutations as prognostic and predictive markers in non–small cell lung cancer. *Journal of Thoracic Oncology,* 8**,** 530-542.
Mata-Greenwood, E., Cuendet, M., Sher, D., Gustin, D., Stock, W. & Pezzuto, J. M. 2002. Brusatol-mediated induction of leukemic cell differentiation and g1 arrest is associated with down-regulation of c-myc. *Leukemia,* 16**,** 2275.
Mayeux, R. 2004. Biomarkers: Potential uses and limitations. *NeuroRx,* 1**,** 182-188.
McCall, M., M Bolstad, B. & A Irizarry, R. 2010. *Frozen robust multi-array analysis (frma)*.
McCay, P. B. 1985. Vitamin e: Interactions with free radicals and ascorbate. *Annual Review of Nutrition,* 5**,** 323-340.
McCubrey, J. A., Steelman, L. S., Bertrand, F. E., Davis, N. M., Sokolosky, M., Abrams, S. L., Montalto, G., D'Assoro, A. B., Libra, M., Nicoletti, F., Maestro, R., Basecke, J., Rakus, D., Gizak, A., Demidenko, Z., Cocco, L., Martelli, A. M. & Cervello, M. 2014. Gsk-3 as potential target for therapeutic intervention in cancer. *Oncotarget,* 5**,** 2881-2911.
McMahon, M., Itoh, K., Yamamoto, M. & Hayes, J. 2003. Keap1-dependent proteasomal degradation of transcription factor nrf2 contributes to the negative regulation of antioxidant response element-driven gene expression. *The Journal of biological chemistry,* 278**,** 21592-600.
McMahon, M., Lamont, D. J., Beattie, K. A. & Hayes, J. D. 2010. Keap1 perceives stress via three sensors for the endogenous signaling molecules nitric oxide, zinc, and alkenals. *Proceedings of the National Academy of Sciences of the United States of America,* 107**,** 18838-18843.
McMahon, M., Thomas, N., Itoh, K., Yamamoto, M. & Hayes, J. 2006. Dimerization of substrate adaptors can facilitate cullin-mediated ubiquitylation of proteins by a “tethering” mechanism □ s. *Journal of Biological Chemistry,* 281**,** 24756-68.
McMahon, M., Thomas, N., Itoh, K., Yamamoto, M. & Hayes, J. D. 2004. Redox-regulated turnover of nrf2 is determined by at least two separate protein domains, the redox-sensitive neh2 degron and the redox-insensitive neh6 degron. *Journal of Biological Chemistry,* 279**,** 31556-31567.

Meister, A. 1983. Selective modification of glutathione metabolism. *Science,* 220**,** 472-477.
Meister, A. & Anderson, M., E. 1983. Glutathione. *Annual Review of Biochemistry,* 52**,** 711-760.
Meitzler, J., Antony, S., Wu, Y., Juhasz, A., Liu, H., Jiang, G., Lu, J., Roy, K. & H Doroshow, J. 2013. Nadph oxidases: A perspective on reactive oxygen species production in tumor biology. *Antioxidant Redox Signalling* 20**,** 2873-2889.
Metz, T., Harris, A. W. & Adams, J. M. 1995. Absence of p53 allows direct immortalization of hematopoietic cells by the myc and raf oncogenes. *Cell,* 82**,** 29-36.
Metzger, M. B., Hristova, V. A. & Weissman, A. M. 2012. Hect and ring finger families of e3 ubiquitin ligases at a glance. *Journal of Cell Science,* 125**,** 531.
Miao, W., Hu, L., James Scrivens, P. & Batist, G. 2005. Transcriptional regulation of nf-e2 p45-related factor (nrf2) expression by the aryl hydrocarbon receptor-xenobiotic response element signaling pathway - direct cross-talk between phase i and ii drug-metabolizing enzymes. *The Journal of biological chemistry,* 280**,** 20340-8.
Miao, W., Hu, L., Kandouz, M., Hamilton, D. & Batist, G. 2004. A cell-based system to identify and characterize the molecular mechanism of drug-metabolizing enzyme (dme) modulators. *Biochemical Pharmacology,* 67**,** 1897-1905.
Miller, C. J., Gounder, S. S., Kannan, S., Goutam, K., Muthusamy, V. R., Firpo, M. A., Symons, J. D., Paine, R., Hoidal, J. R. & Rajasekaran, N. S. 2012. Disruption of nrf2/are signaling impairs antioxidant mechanisms and promotes cell degradation pathways in aged skeletal muscle. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease,* 1822**,** 1038-1050.
Mitsuishi, Y., Taguchi, K., Kawatani, Y., Shibata, T., Nukiwa, T., Aburatani, H., Yamamoto, M. & Motohashi, H. 2012. Nrf2 redirects glucose and glutamine into anabolic pathways in metabolic reprogramming. *Cancer Cell,* 22**,** 66-79.
Mo, C., Wang, L., Zhang, J., Numazawa, S., Tang, H., Tang, X., Han, X., Li, J., Yang, M., Wang, Z., Wei, D. & Xiao, H. 2014. The crosstalk between nrf2 and ampk signal pathways is important for the anti-inflammatory effect of berberine in lps-stimulated macrophages and endotoxin-shocked mice. *Antioxidants & Redox Signaling,* 20**,** 574-588.
Mohler, J., Vani, K., Leung, S. & Epstein, A. 1991. Segmentally restricted, cephalic expression of a leucine zipper gene during drosophila embryogenesis. *Mechanisms of Development,* 34**,** 3-9.
Moi, P., Chan, K., Asunis, I., Cao, A. & Kan, Y. W. 1994. Isolation of nf-e2-related factor 2 (nrf2), a nf-e2-like basic leucine zipper transcriptional activator that binds to the tandem nf-e2/ap1 repeat of the beta-globin locus control region. *Proceedings of the National Academy of Sciences of the United States of America,* 91**,** 9926-9930.
Morales-Pantoja, I. E., Hu, C.-l., Perrone-Bizzozero, N. I., Zheng, J. & Bizzozero, O. A. 2016. Nrf2-dysregulation correlates with reduced synthesis and low glutathione levels in experimental autoimmune encephalomyelitis. *Journal of neurochemistry,* 139**,** 640-650.
Moreira, S., Fonseca, I., Nunes, M. J., Rosa, A., Lemos, L., Rodrigues, E., Carvalho, A. N., Outeiro, T. F., Rodrigues, C. M. P., Gama, M. J. & Castro-Caldas, M. 2017. Nrf2 activation by tauroursodeoxycholic acid in experimental models of parkinson's disease. *Experimental Neurology,* 295**,** 77-87.
Morgan, M. J. & Liu, Z.-g. 2010. Crosstalk of reactive oxygen species and nf-κb signaling. *Cell Research,* 21**,** 103.
Motohashi, H., Katsuoka, F., Engel, J. D. & Yamamoto, M. 2004. Small maf proteins serve as transcriptional cofactors for keratinocyte differentiation in the keap1–nrf2 regulatory pathway. *Proceedings of the National Academy of Sciences of the United States of America,* 101**,** 6379-6384.
Mustacich, D. & Powis, G. 2000. Thioredoxin reductase. *Biochemical Journal,* 346**,** 1-8.
Nanavaty, P., S Alvarez, M. & Michael Alberts, W. 2014. Lung cancer screening: Advantages, controversies, and applications. *Cancer control : journal of the Moffitt Cancer Center***,** 9-14.
Newsholme, P., Cruzat, V. F., Keane, K. N., Carlessi, R. & de Bittencourt, P. I. H. 2016. Molecular mechanisms of ros production and oxidative stress in diabetes. *Biochemical Journal,* 473**,** 4527-4550.
Nickel, A., Kohlhaas, M. & Maack, C. 2014. Mitochondrial reactive oxygen species production and elimination. *Journal of Molecular and Cellular Cardiology,* 73**,** 26-33.
Nioi, P., McMahon, M., Itoh, K., Yamamoto, M. & Hayes, J. D. 2003. Identification of a novel nrf2-regulated antioxidant response element (are) in the mouse nad(p)h:Quinone oxidoreductase 1 gene: Reassessment of the are consensus sequence. *Biochemical Journal,* 374**,** 337-348.

Nioi, P., Nguyen, T., Sherratt, P. J. & Pickett, C. B. 2005. The carboxy-terminal neh3 domain of nrf2 is required for transcriptional activation. *Molecular and Cellular Biology,* 25**,** 10895-10906.

Nishida, N., Yano, H., Nishida, T., Kamura, T. & Kojiro, M. 2006. Angiogenesis in cancer. *Vascular Health Risk Management,* 2**,** 213-9.

Nogueira, V. & Hay, N. 2013. Molecular pathways: Reactive oxygen species homeostasis in cancer cells and implications for cancer therapy. *Clinical Cancer Research,* 19**,** 4309-4314.

Noll, C., Tlili, A., Ripoll, C., Mallet, L., Paul, J.-L., Delabar, J.-M. & Janel, N. 2012. Dyrk1a activates antioxidant nqo1 expression through an erk1/2–nrf2 dependent mechanism. *Molecular Genetics and Metabolism,* 105**,** 484-488.

Nouhi, Z., Chevillard, G., Derjuga, A. & Blank, V. 2007. Endoplasmic reticulum association and n-linked glycosylation of the human nrf3 transcription factor. *FEBS Letters,* 581**,** 5401-5406.

Ohta, T., Iijima, K., Miyamoto, M., Nakahara, I., Tanaka, H., Ohtsuji, M., Suzuki, T., Kobayashi, A., Yokota, J., Sakiyama, T., Shibata, T., Yamamoto, M. & Hirohashi, S. 2008. Loss of keap1 function activates nrf2 and provides advantages for lung cancer cell growth. *Cancer research,* 68**,** 1303-9.

Ohtsuji, M., Katsuoka, F., Kobayashi, A., Aburatani, H., Hayes, J. & Yamamoto, M. 2008. Nrf1 and nrf2 play distinct roles in activation of antioxidant response element-dependent genes. *The Journal of biological chemistry,* 283**,** 33554-62.

Olayanju, A., Copple, I. M., Bryan, H. K., Edge, G. T., Sison, R. L., Wong, M. W., Lai, Z.-Q., Lin, Z.-X., Dunn, K., Sanderson, C. M., Alghanem, A. F., Cross, M. J., Ellis, E. C., Ingelman-Sundberg, M., Malik, H. Z., Kitteringham, N. R., Goldring, C. E. & Park, B. K. 2015. Brusatol provokes a rapid and transient inhibition of nrf2 signaling and sensitizes mammalian cells to chemical toxicity—implications for therapeutic targeting of nrf2. *Free Radical Biology and Medicine,* 78**,** 202-212.

Ooi, A., Dykema, K., Ansari, A., Petillo, D., Snider, J., Kahnoski, R., Anema, J., Craig, D., Carpten, J., Teh, B.-T. & Furge, K. A. 2013. cul3 and nrf2 mutations confer an nrf2 activation phenotype in a sporadic form of papillary renal cell carcinoma. *Cancer Research,* 73**,** 2044-2051.

Ooi, A., Wong, J.-C., Petillo, D., Roossien, D., Perrier-Trudova, V., Whitten, D., Min, Bernice Wong H., Tan, M.-H., Zhang, Z., Yang, Ximing J., Zhou, M., Gardie, B., Molinié, V., Richard, S., Tan, Puay H., Teh, Bin T. & Furge, Kyle A. 2011. An antioxidant response phenotype shared between hereditary and sporadic type 2 papillary renal cell carcinoma. *Cancer Cell,* 20**,** 511-523.

Padmanabhan, B., Tong, K. I., Ohta, T., Nakamura, Y., Scharlock, M., Ohtsuji, M., Kang, M.-I., Kobayashi, A., Yokoyama, S. & Yamamoto, M. 2006. Structural basis for defects of keap1 activity provoked by its point mutations in lung cancer. *Molecular Cell,* 21**,** 689-700.

Patel, S., Doble, B. W., MacAulay, K., Sinclair, E. M., Drucker, D. J. & Woodgett, J. R. 2008. Tissue-specific role of glycogen synthase kinase 3β in glucose homeostasis and insulin action. *Molecular and Cellular Biology,* 28**,** 6314-6328.

Patel, S. & Woodgett, J. I. M. 2008. Glycogen synthase kinase-3 and cancer: Good cop, bad cop? *Cancer cell,* 14**,** 351-353.

Peng H, Li H, Sheehy A, Cullen P, Allaire N, Scannevin RH. 2016. Dimethyl fumarate alters microglia phenotype and protects neurons against proinflammatory toxic microenvironments. *Journal of Neuroimmunology,* 299,35-44.

Penning, T. M. 2017. Aldo-keto reductase regulation by the nrf2 system: Implications for stress response, chemotherapy drug resistance, and carcinogenesis. *Chemical Research in Toxicology,* 30**,** 162-176.

Pfaffl, M. W. 2001. A new mathematical model for relative quantification in real-time rt–pcr. *Nucleic Acids Research,* 29**,** 45.

Pickart, C. M. 2001. Mechanisms underlying ubiquitination. *Annual Review of Biochemistry,* 70**,** 503-533.

Pickering, A., Vojtovich, L., Tower, J. & J A Davies, K. 2012. Oxidative stress adaptation with acute, chronic and repeated stress. *Free Radical Biology & Medicine,* 55**,** 109-118.

Pierce, G. B., Parchment, R. E. & Lewellyn, A. L. 1991. Hydrogen peroxide as a mediator of programmed cell death in the blastocyst. *Differentiation,* 46**,** 181-186.

Pikor, L. A., Ramnarine, V. R., Lam, S. & Lam, W. L. 2013. Genetic alterations defining nsclc subtypes and their therapeutic implications. *Lung Cancer,* 82**,** 179-189.

Pintard, L., Willems, A. & Peter, M. 2004. Cullin-based ubiquitin ligases: Cul3–btb complexes join the family. *The EMBO Journal,* 23**,** 1681-1687.

Poirier, Y., Antonenkov, V. D., Glumoff, T. & Hiltunen, J. K. 2006. Peroxisomal β-oxidation—a metabolic pathway with multiple functions. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research,* 1763**,** 1413-1426.

Pronobis, M. I., Rusan, N. M. & Peifer, M. 2015. A novel gsk3-regulated apc:Axin interaction regulates wnt signaling by driving a catalytic cycle of efficient βcatenin destruction. *eLife,* 4**,** 08022.

Puchalski, R. B. & Fahl, W. E. 1990. Expression of recombinant glutathione s-transferase pi, ya, or yb1 confers resistance to alkylating agents. *Proceedings of the National Academy of Sciences of the United States of America,* 87**,** 2443-2447.

Qiu, C., Kivipelto, M. & von Strauss, E. 2009. Epidemiology of alzheimer's disease: Occurrence, determinants, and strategies toward intervention. *Dialogues in Clinical Neuroscience,* 11**,** 111-128.

Rabbani, P. S., Ellison, T., Waqas, B., Sultan, D., Abdou, S., David, J. A., Cohen, J. M., Gomez-Viso, A., Lam, G., Kim, C., Thomson, J. & Ceradini, D. J. 2018. Targeted nrf2 activation therapy with rta 408 enhances regenerative capacity of diabetic wounds. *Diabetes Research and Clinical Practice,* 139**,** 11-23.

Rada, P., Rojo, A. I., Chowdhry, S., McMahon, M., Hayes, J. D. & Cuadrado, A. 2011. Scf/β-trcp promotes glycogen synthase kinase 3-dependent degradation of the nrf2 transcription factor in a keap1-independent manner. *Molecular and Cellular Biology,* 31**,** 1121-1133.

Rada, P., Rojo, A. I., Evrard-Todeschi, N., Innamorato, N. G., Cotte, A., Jaworski, T., Tobón-Velasco, J. C., Devijver, H., García-Mayoral, M. F., Van Leuven, F., Hayes, J. D., Bertho, G. & Cuadrado, A. 2012. Structural and functional characterization of nrf2 degradation by the glycogen synthase kinase 3/β-trcp axis. *Molecular and Cellular Biology,* 32**,** 3486-3499.

Raghunath, A., Sundarraj, K., Nagarajan, R., Arfuso, F., Bian, J., Kumar, A. P., Sethi, G. & Perumal, E. 2018. Antioxidant response elements: Discovery, classes, regulation and potential applications. *Redox Biology,* 17**,** 297-314.

Rahman, I., Kode, A. & Biswas, S. K. 2007. Assay for quantitative determination of glutathione and glutathione disulfide levels using enzymatic recycling method. *Nature Protocols,* 1**,** 3159.

Ramakrishnan, S., Anand, V. & Roy, S. 2014. Vascular endothelial growth factor signaling in hypoxia and inflammation. *Journal of Neuroimmune Pharmacology,* 9**,** 142-160.

Ramsey, C. P., Glass, C. A., Montgomery, M. B., Lindl, K. A., Ritson, G. P., Chia, L. A., Hamilton, R. L., Chu, C. T. & Jordan-Sciutto, K. L. 2007. Expression of nrf2 in neurodegenerative diseases. *Journal of neuropathology and experimental neurology,* 66**,** 75-85.

Rappa, G., Lorico, A., Flavell, R. & C Sartorelli, A. 1998. Evidence that the multidrug resistance protein (mrp) functions as a co-transporter of glutathione and natural product toxins. *Cancer research,* 57**,** 5232-7.

Reddy, N. M., Kleeberger, S. R., Cho, H.-Y., Yamamoto, M., Kensler, T. W., Biswal, S. & Reddy, S. P. 2007. Deficiency in nrf2-gsh signaling impairs type ii cell growth and enhances sensitivity to oxidants. *American journal of respiratory cell and molecular biology,* 37**,** 3-8.

Ren, D., Villeneuve, N. F., Jiang, T., Wu, T., Lau, A., Toppin, H. A. & Zhang, D. D. 2011. Brusatol enhances the efficacy of chemotherapy by inhibiting the nrf2-mediated defense mechanism. *Proceedings of the National Academy of Sciences,* 108**,** 1433-1438.

Robertson, H., Hayes, J. D. & Sutherland, C. 2018. A partnership with the proteasome; the destructive nature of gsk3. *Biochemical Pharmacology,* 147**,** 77-92.

Rojo, A. I., Rada, P., Egea, J., Rosa, A. O., López, M. G. & Cuadrado, A. 2008. Functional interference between glycogen synthase kinase-3 beta and the transcription factor nrf2 in protection against kainate-induced hippocampal celldeath. *Molecular and Cellular Neuroscience,* 39**,** 125-132.

Rojo Ana, I., Sagarra María Rosa, d. & Cuadrado, A. 2008. Gsk-3β down-regulates the transcription factor nrf2 after oxidant damage: Relevance to exposure of neuronal cells to oxidative stress. *Journal of Neurochemistry,* 105**,** 192-202.

Rojo de la Vega, M., Chapman, E. & Zhang, D. D. 2018. Nrf2 and the hallmarks of cancer. *Cancer Cell,* 34**,** 21-43.

Romero, R., Sayin, V. I., Davidson, S. M., Bauer, M. R., Singh, S. X., LeBoeuf, S. E., Karakousi, T. R., Ellis, D. C., Bhutkar, A., Sanchez-Rivera, F. J., Subbaraj, L., Martinez, B., Bronson, R. T., Prigge, J. R., Schmidt, E. E., Thomas, C. J., Goparaju, C., Davies, A.,

Dolgalev, I., Heguy, A., Allaj, V., Poirier, J. T., Moreira, A. L., Rudin, C. M., Pass, H. I., Vander Heiden, M. G., Jacks, T. & Papagiannakopoulos, T. 2017. Keap1 loss promotes kras-driven lung cancer and results in a dependence on glutaminolysis. *Nature medicine,* 23**,** 1362-1368.

Roomi, M., Monterrey, J. C., Kalinovsky, T., Rath, M. & Niedzwiecki, A. 2009. Patterns of mmp-2 and mmp-9 expression in human cancer cell lines. *Oncology Reports,* 21**,** 1323-33.

Rotblat, B., Melino, G. & A Knight, R. 2012. Nrf2 and p53: Januses in cancer? *Oncotarget,* 3**,** 1272-1283.

Rudolph J. 2005. Redox Regulation of the Cdc25 Phosphatases. *Antioxidants & Redox Signaling*, 7(5-6), 761-7.

Ruscoe, J. E., Rosario, L. A., Wang, T., Gaté, L., Arifoglu, P., Wolf, C. R., Henderson, C. J., Ronai, Z. e. & Tew, K. D. 2001. Pharmacologic or genetic manipulation of glutathione s-transferase p1–1 (gstπ) influences cell proliferation pathways. *Journal of Pharmacology and Experimental Therapeutics,* 298**,** 339-345.

Rushmore, T., H., Morton, M. & Pickett, C. 1991. The anti-oxidative stress and identification of the DNA consensus sequence required for functional activity. *The Journal of biological chemistry,* 266**,** 11632-9.

Rushmore, T. H. & Pickett, C. 1990. Transcriptional regulation of the rat glutathione s-transferase ya subunit gene. Characterization of a xenobiotic-responsive element controlling inducible expression by phenolic antioxidants. *The Journal of biological chemistry,* 265**,** 14648-53.

Sadeghi, M. R., Jeddi, F., Soozangar, N., Somi, M. H. & Samadi, N. 2017. The role of nrf2-keap1 axis in colorectal cancer, progression, and chemoresistance. *Tumor Biology,* 39**,** 1010428317705510.

Saito, R., Suzuki, T., Hiramoto, K., Asami, S., Naganuma, E., Suda, H., Iso, T., Yamamoto, H., Morita, M., Baird, L., Furusawa, Y., Negishi, T., Ichinose, M. & Yamamoto, M. 2016. Characterizations of three major cysteine sensors of keap1 in stress response. *Molecular and Cellular Biology,* 36**,** 271-284.

Saitoh, M., Nishitoh, H., Fujii, M., Takeda, K., Tobiume, K., Sawada, Y., Kawabata, M., Miyazono, K. & Ichijo, H. 1998. Mammalian thioredoxin is a direct inhibitor of apoptosis signal-regulating kinase (ask) 1. *The EMBO Journal,* 17**,** 2596-2606.

Salazar-Roa, M., I Rojo, A., Velasco, D., Sagarra, M. & Cuadrado, A. 2006. Glycogen synthase kinase-3?? Inhibits the xenobiotic and antioxidant cell response by direct phosphorylation and nuclear exclusion of the transcription factor nrf2. *The Journal of biological chemistry,* 281**,** 14841-51.

Samet, J. M., Avila-Tang, E., Boffetta, P., Hannan, L. M., Olivo-Marston, S., Thun, M. J. & Rudin, C. M. 2009. Lung cancer in never smokers: Clinical epidemiology and environmental risk factors. *Clinical cancer research : an official journal of the American Association for Cancer Research,* 15**,** 5626-5645.

Sarbassov, D. D., Guertin, D. A., Ali, S. M. & Sabatini, D. M. 2005. Phosphorylation and regulation of akt/pkb by the rictor-mtor complex. *Science,* 307**,** 1098-1101.

Sasaki, H., Hikosaka, Y., Okuda, K., Kawano, O., Moriyama, S., Yano, M. & Fujii, Y. 2010. nfe2l2 gene mutation in male japanese squamous cell carcinoma of the lung. *Journal of Thoracic Oncology,* 5**,** 786-789.

Sasaki, H., Sato, H., Kuriyama-Matsumura, K., Sato, K., Maebara, K., Wang, H., Tamba, M., Itoh, K., Yamamoto, M. & Bannai, S. 2002. Electrophile response element-mediated induction of the cystine/glutamate exchange transporter gene expression. *The Journal of biological chemistry,* 277**,** 44765-71.

Sasaki, H., Shitara, M., Yokota, K., Hikosaka, Y., Moriyama, S., Yano, M. & Fujii, Y. 2012. Increased nrf2 gene (nfe2l2) copy number correlates with mutations in lung squamous cell carcinomas. *Molecular medicine reports,* 6**,** 391-4.

Sasaki, H., Suzuki, A., Shitara, M., Hikosaka, Y., Okuda, K., Moriyama, S., Yano, M. & Fujii, Y. 2013a. Genotype analysis of the nrf2 gene mutation in lung cancer. *International journal of molecular medicine,* 31**,** 1135-1138.

Sasaki, H., Suzuki, A., Shitara, M., Okuda, K., Hikosaka, Y., Moriyama, S., Yano, M. & Fujii, Y. 2013b. Keap1 mutations in lung cancer patients. *Oncology letters,* 6**,** 719-721.

Scheffler, M., Bos, M., Gardizi, M., König, K., Michels, S., Fassunke, J., Heydt, C., Pasternack, H., Ihle, M., Ueckeroth, F., Albus, K., Serke, M., Gerigk, U., Schulte, W., Töpelt, K., Nogova, L., Zander, T., Engel-Riedel, W., Stoelben, E. & Wolf, J. 2014. Pik3ca mutations in non-small cell lung cancer (nsclc): Genetic heterogeneity, prognostic impact and incidence of prior malignancies. *Oncotarget,* 6**,** 1315-1326.

Schieber M, Chandel NS. 2013. ROS function in redox signaling and oxidative stress. *Current biol*ogy, 24(10): 453-62.

Schnelldorfer, T., Gansauge, S., Gansauge, F., Schlosser, S., Beger Hans, G. & Nussler Andreas, K. 2000. Glutathione depletion causes cell growth inhibition and enhanced apoptosis in pancreatic cancer cells. *Cancer,* 89**,** 1440-1447.

Sekine, H., Okazaki, K., Ota, N., Shima, H., Katoh, Y., Suzuki, N., Igarashi, K., Ito, M., Motohashi, H. & Yamamoto, M. 2015. The mediator subunit med16 transduces nrf2-activating signals into antioxidant gene expression. *Molecular and cellular biology,* 36**,** 407-420.

Sekine, H., Okazaki, K., Ota, N., Shima, H., Katoh, Y., Suzuki, N., Igarashi, K., Ito, M., Motohashi, H. & Yamamoto, M. 2016. The mediator subunit med16 transduces nrf2-activating signals into antioxidant gene expression. *Molecular and Cellular Biology,* 36**,** 407-420.

Seo, E., Kim, H., Kim, R., Yun, S., Kim, M., Han, J.-K., Costantini, F. & Jho, E.-h. 2009. Multiple isoforms of β-trcp display differential activities in the regulation of wnt signaling. *Cellular Signalling,* 21**,** 43-51.

Sergi, A., Ariadna, L. & Susana de la, L. 2011. Dyrk family of protein kinases: Evolutionary relationships, biochemical properties, and functional roles. *The FASEB Journal,* 25**,** 449-462.

Shah, S. A., Potter, M. W. & Callery, M. P. 2001. Ubiquitin proteasome pathway: Implications and advances in cancer therapy. *Surgical Oncology,* 10**,** 43-52.

Shibata, T., Ohta, T., Tong, K. I., Kokubu, A., Odogawa, R., Tsuta, K., Asamura, H., Yamamoto, M. & Hirohashi, S. 2008. Cancer related mutations in nrf2 impair its recognition by keap1-cul3 e3 ligase and promote malignancy. *Proceedings of the National Academy of Sciences of the United States of America,* 105**,** 13568-13573.

Shin, S.-K., Cho, H.-W., Song, S.-E. & Song, D.-K. 2018. Catalase and nonalcoholic fatty liver disease. *Pflügers Archiv - European Journal of Physiology*.

Shiono, S., Katahira, M., Abiko, M. & Sato, T. 2015. Smoking is a perioperative risk factor and prognostic factor for lung cancer surgery. *General Thoracic and Cardiovascular Surgery,* 63**,** 93-98.

Shivdasani, R. A., Rosenblatt, M. F., Zucker-Franklin, D., Jackson, C. W., Hunt, P., Saris, C. J. M. & Orkin, S. H. 1995. Transcription factor nf-e2 is required for platelet formation independent of the actions of thrombopoeitin/mgdf in megakaryocyte development. *Cell,* 81**,** 695-704.

Siddik, Z. H. 2002. Cisplatin: Mode of cytotoxic action and molecular basis of resistance. *Oncogene,* 22**,** 7265.

Simon, H. U., Haj-Yehia, A. & Levi-Schaffer, F. 2000. Role of reactive oxygen species (ros) in apoptosis induction. *Apoptosis,* 5**,** 415-418.

Singh, A., Misra, V., Thimmulappa, R. K., Lee, H., Ames, S., Hoque, M. O., Herman, J. G., Baylin, S. B., Sidransky, D., Gabrielson, E., Brock, M. V. & Biswal, S. 2006. Dysfunctional keap1–nrf2 interaction in non-small-cell lung cancer. *PLOS Medicine,* 3**,** 420.

Singh, A., Rangasamy, T., Thimmulappa, R., Lee, H., Osburn, W., Brigelius-Flohé, R., W Kensler, T., Yamamoto, M. & Biswal, S. 2007. Glutathione peroxidase 2, the major cigarette smoke–inducible isoform of gpx in lungs, is regulated by nrf2. *American journal of respiratory cell and molecular biology,* 35**,** 639-50.

Singh, A., Venkannagari, S., H Oh, K., Zhang, Y.-Q., Liu, L., Rohde, J., Nimmagadda, S., Sudini, K., Brimacombe, K., Gajghate, S., Ma, J., Wang, A., Xu, X., Shahane, S., Xia, M., Woo, J., Acquaah-Mensah, G., Wang, Z., Ferrer, M. & Biswal, S. 2016. Small molecule inhibitor of nrf2 selectively intervenes therapeutic resistance in keap1-deficient nsclc tumors. *ASC Chemical Biology (ACS Publications),* 11**,** 3214-3225.

Skibinski, G., Hwang, V., Ando, D. M., Daub, A., Lee, A. K., Ravisankar, A., Modan, S., Finucane, M. M., Shaby, B. A. & Finkbeiner, S. 2017. Nrf2 mitigates lrrk2- and α-synuclein–induced neurodegeneration by modulating proteostasis. *Proceedings of the National Academy of Sciences of the United States of America,* 114**,** 1165-1170.

Soga, M., Matsuzawa, A. & Ichijo, H. 2012. Oxidative stress-induced diseases via the ask1 signaling pathway. *International Journal of Cell Biology,* 2012**,** 439587.

Soppa, U. & Becker, W. 2015. Dyrk protein kinases. *Current Biology,* 25**,** R488-R489.

Spataro, V., Norbury, C. & Harris, A. L. 1998. The ubiquitin-proteasome pathway in cancer. *British Journal of Cancer,* 77**,** 448-455.

Sporn, M. B. & Liby, K. T. 2012. Nrf2 and cancer: The good, the bad and the importance of context. *Nature Reviews Cancer,* 12**,** 564.

Squadrito, G. L. & Pryor, W. A. 1995. The formation of peroxynitrite in vivo from nitric oxide and superoxide. *Chemico-Biological Interactions,* 96**,** 203-206.

Storz, P. 2016. Kras, ros and the initiation of pancreatic cancer. *Small GTPases,* 8**,** 38-42.

Sun, Z., Chin, Y. E. & Zhang, D. D. 2009. Acetylation of nrf2 by p300/cbp augments promoter-specific DNA binding of nrf2 during the antioxidant response. *Molecular and Cellular Biology,* 29**,** 2658-2672.

Sunaga N, Shames DS, Girard L, Peyton M, Larsen JE, Imai H, et al. 2011. Knockdown of oncogenic KRAS in non-small cell lung cancers suppresses tumor growth and sensitizes tumor cells to targeted therapy. *Molecular cancer therapeutics,*10(2), 336-46.

Sutherland, C. 2011. What are the bona fide gsk3 substrates? *International Journal of Alzheimer's Disease,* 2011**,** 505607.

Swatek, K. N. & Komander, D. 2016. Ubiquitin modifications. *Cell Research,* 26**,** 399.

Taguchi, K., Fujikawa, N., Komatsu, M., Ishii, T., Unno, M., Akaike, T., Motohashi, H. & Yamamoto, M. 2012. Keap1 degradation by autophagy for the maintenance of redox homeostasis. *Proceedings of the National Academy of Sciences,* 109**,** 13561-13566.

Taguchi, K. & Yamamoto, M. 2017. The keap1–nrf2 system in cancer. *Frontiers in Oncology,* 7**,** 85.

Takaya, K., Suzuki, T., Motohashi, H., Onodera, K., Satomi, S., Kensler, T. W. & Yamamoto, M. 2012. Validation of the multiple sensor mechanism of the keap1-nrf2 system. *Free radical biology & medicine,* 53**,** 817-827.

Tao, S., Wang, S., Moghaddam, S. J., Ooi, A., Chapman, E., Wong, P. K. & Zhang, D. D. 2014. Oncogenic kras confers chemoresistance by upregulating nrf2. *Cancer research,* 74**,** 7430-7441.

Tarver, H. & Schmtdt, C. L. A. 1939. The conversion of methionine to cystine: Experiments with radioactive sulfur. *Journal of Biological Chemistry,* 130**,** 67-80.

Tebay, L. E., Robertson, H., Durant, S. T., Vitale, S. R., Penning, T. M., Dinkova-Kostova, A. T. & Hayes, J. D. 2015. Mechanisms of activation of the transcription factor nrf2 by redox stressors, nutrient cues, and energy status and the pathways through which it attenuates degenerative disease. *Free Radical Biology and Medicine,* 88**,** 108-146.

Tew, K. D. 1994. Glutathione-associated enzymes in anticancer drug resistance. *Cancer Research,* 54**,** 4313-4320.

The Cancer Genome Atlas, N., Lawrence, M. S., Sougnez, C., Lichtenstein, L., Cibulskis, K., Lander, E., Gabriel, S. B., Getz, G., Ally, A., Balasundaram, M., Birol, I., Bowlby, R., Brooks, D., Butterfield, Y. S. N., Carlsen, R., Cheng, D., Chu, A., Dhalla, N., Guin, R., Holt, R. A., Jones, S. J. M., Lee, D., Li, H. I., Marra, M. A., Mayo, M., Moore, R. A., Mungall, A. J., Gordon Robertson, A., Schein, J. E., Sipahimalani, P., Tam, A., Thiessen, N., Wong, T., Protopopov, A., Santoso, N., Lee, S., Parfenov, M., Zhang, J., Mahadeshwar, H. S., Tang, J., Ren, X., Seth, S., Haseley, P., Zeng, D., Yang, L., Xu, A. W., Song, X., Pantazi, A., Bristow, C. A., Hadjipanayis, A., Seidman, J., Chin, L., Park, P. J., Kucherlapati, R., Akbani, R., Casasent, T., Liu, W., Lu, Y., Mills, G., Motter, T., Weinstein, J., Diao, L., Wang, J., Hong Fan, Y., Liu, J., Wang, K., Todd Auman, J., Balu, S., Bodenheimer, T., Buda, E., Neil Hayes, D., Hoadley, K. A., Hoyle, A. P., Jefferys, S. R., Jones, C. D., Kimes, P. K., Liu, Y., Marron, J. S., Meng, S., Mieczkowski, P. A., Mose, L. E., Parker, J. S., Perou, C. M., Prins, J. F., Roach, J., Shi, Y., Simons, J. V., Singh, D., Soloway, M. G., Tan, D., Veluvolu, U., Walter, V., Waring, S., Wilkerson, M. D., Wu, J., Zhao, N., Cherniack, A. D., Hammerman, P. S., Tward, A. D., Sekhar Pedamallu, C., et al. 2015. Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature,* 517**,** 576.

The Cancer Genome Atlas Research, N. 2012. Comprehensive genomic characterization of squamous cell lung cancers. *Nature,* 489**,** 519.

The Cancer Genome Atlas Research, N. 2014. Comprehensive molecular profiling of lung adenocarcinoma. *Nature,* 511**,** 543.

Thimmulappa, R. K., Mai, K. H., Srisuma, S., Kensler, T. W., Yamamoto, M. & Biswal, S. 2002. Identification of nrf2-regulated genes induced by the chemopreventive agent sulforaphane by oligonucleotide microarray. *Cancer Research,* 62**,** 5196-5203.

Tian, Y., Wu, K., Liu, Q., Han, N., Zhang, L., Chu, Q. & Chen, Y. 2016. Modification of platinum sensitivity by keap1/nrf2 signals in non-small cell lung cancer. *Journal of Hematology & Oncology,* 9**,** 83.

Tod, A. M., Craven, J. & Allmark, P. 2008. Diagnostic delay in lung cancer: A qualitative study. *Journal of Advanced Nursing,* 61**,** 336-343.

Todorovic, M., Wood, S. A. & Mellick, G. D. 2016. Nrf2: A modulator of parkinson's disease? *Journal of Neural Transmission,* 123**,** 611-619.

Toki, T., Itoh, J., Kitazawa, J. i., Arai, K., Hatakeyama, K., Akasaka, J.-i., Igarashi, K., Nomura, N., Yokoyama, M., Yamamoto, M. & Ito, E. 1997. Human small maf proteins form heterodimers with cnc family transcription factors and recognize the nf-e2 motif. *Oncogene,* 14**,** 1901-10.

Tong, K. I., Katoh, Y., Kusunoki, H., Itoh, K., Tanaka, T. & Yamamoto, M. 2006. Keap1 recruits neh2 through binding to etge and dlg motifs: Characterization of the two-site molecular recognition model. *Molecular and Cellular Biology,* 26**,** 2887-2900.

Tong Kit, I., Kobayashi, A., Katsuoka, F. & Yamamoto, M. 2006. Two-site substrate recognition model for the keap1-nrf2 system: A hinge and latch mechanism. *Biological Chemistry,* 387**,** 1311.

Torre, L., Siegel, R. & Ahmedin, J. 2015. Lung cancer and personalized medicine : Current knowledge and therapies. *In:* AHMAD, A. & GADGEEL, S. (eds.) *Advances in experimental medicine and biology.*

Townsend, D., M & Tew, K. 2003. *Townsend, d. M. & tew, k. D. The role of glutathione-s-transferase in anti-cancer drug resistance. Oncogene 22, 7369-7375*.

Traber, M. G. & Atkinson, J. 2007. Vitamin e, antioxidant and nothing more. *Free radical biology & medicine,* 43**,** 4-15.

Traber, M. G. & Stevens, J. F. 2011. Vitamins c and e: Beneficial effects from a mechanistic perspective. *Free radical biology & medicine,* 51**,** 1000-1013.

Traverso, N., Ricciarelli, R., Nitti, M., Marengo, B., Furfaro, A. L., Adelaide Pronzato, M., Marinari, U. & Domenicotti, C. 2013. Role of glutathione in cancer progression and chemoresistance. *Oxidative medicine and cellular longevity,* 2013**,** 10.

Tsuchiya, Y., Morita, T., Kim, M., Iemura, S.-i., Natsume, T., Yamamoto, M. & Kobayashi, A. 2011. Dual regulation of the transcriptional activity of nrf1 by β-trcp- and hrd1-dependent degradation mechanisms. *Molecular and Cellular Biology,* 31**,** 4500-4512.

Tsuchiya, Y., Taniguchi, H., Ito, Y., Morita, T., Karim, M. R., Ohtake, N., Fukagai, K., Ito, T., Okamuro, S., Iemura, S.-i., Natsume, T., Nishida, E. & Kobayashi, A. 2013. The ck2-nrf1 axis controls the clearance of ubiquitinated proteins by regulating proteasome gene expression. *Molecular and Cellular Biology,* 33**,** 3461-72.

Tsujita, T., Peirce, V., Baird, L., Matsuyama, Y., Takaku, M., Walsh, S. V., Griffin, J. L., Uruno, A., Yamamoto, M. & Hayes, J. D. 2014. Transcription factor nrf1 negatively regulates the cystine/glutamate transporter and lipid-metabolizing enzymes. *Molecular and Cellular Biology,* 34**,** 3800-3816.

Tung, M.-C., Lin, P.-L., Wang, Y.-C., He, T.-Y., Lee, M.-C., Yeh, S.-D., Tai, C.-J. & Lee, H. 2015. Mutant p53 confers chemoresistance in non-small cell lung cancer by upregulating nrf2. *Oncotarget,* 6**,** 41692-702.

Turrens, J. F. 2003. Mitochondrial formation of reactive oxygen species. *The Journal of Physiology,* 552**,** 335-344.

Ugolkov, A., Gaisina, I., Zhang, J.-S., Billadeau, D. D., White, K., Kozikowski, A., Jain, S., Cristofanilli, M., Giles, F., O'Halloran, T., Cryns, V. L. & Mazar, A. P. 2016. Gsk-3 inhibition overcomes chemoresistance in human breast cancer. *Cancer Letters,* 380**,** 384-392.

Uruno, A., Furusawa, Y., Yagishita, Y., Fukutomi, T., Muramatsu, H., Negishi, T., Sugawara, A., Kensler, T. W. & Yamamoto, M. 2013. The keap1-nrf2 system prevents onset of diabetes mellitus. *Molecular and Cellular Biology,* 33**,** 2996-3010.

Vander Heiden, M. G., Cantley, L. C. & Thompson, C. B. 2009. Understanding the warburg effect: The metabolic requirements of cell proliferation. *Science (New York, N.Y.),* 324**,** 1029-1033.

Varjosalo, M., Keskitalo, S., Van Drogen, A., Nurkkala, H., Vichalkovski, A., Aebersold, R. & Gstaiger, M. 2013. The protein interaction landscape of the human cmgc kinase group. *Cell Reports,* 3**,** 1306-1320.

Verbon, E. H., Post, J. A. & Boonstra, J. 2012. The influence of reactive oxygen species on cell cycle progression in mammalian cells. *Gene,* 511**,** 1-6.

Verschoor, M. L. & Singh, G. 2013. Ets-1 regulates intracellular glutathione levels: Key target for resistant ovarian cancer. *Molecular Cancer,* 12**,** 138.

Viña, J., Borrás, C., Gambini, J., Sastre, J. & Pallardó, F. V. 2005. Why females live longer than males? Importance of the upregulation of longevity-associated genes by oestrogenic compounds. *FEBS Letters,* 579**,** 2541-2545.

Vincent, M. D., Kuruvilla, M. S., Leighl, N. B. & Kamel–Reid, S. 2012. Biomarkers that currently affect clinical practice: Egfr, alk, met, kras. *Current Oncology,* 19**,** S33-S44.

Vogelstein, B., Papadopoulos, N., Velculescu, V. E., Zhou, S., Diaz, L. A. & Kinzler, K. W. 2013. Cancer genome landscapes. *Science (New York, N.Y.),* 339**,** 1546-1558.

Volpe, C. M. O., Villar-Delfino, P. H., dos Anjos, P. M. F. & Nogueira-Machado, J. A. 2018. Cellular death, reactive oxygen species (ros) and diabetic complications. *Cell Death & Disease,* 9**,** 119.

Wang, B. & Williamson, G. 1996. Transcriptional regulation of the human nad(p) h:Quinone oxidoreductase (nqo1) gene by monofunctional inducers. *Biochimica et biophysica acta,* 1307**,** 104-10.

Wang, H., Liu, K., Geng, M., Gao, P., Wu, X., Hai, Y., Li, Y., Li, Y., Luo, L., Hayes, J. D., Wang, X. J. & Tang, X. 2013. Rxrα inhibits the nrf2-are signaling pathway through a direct interaction with the neh7 domain of nrf2. *Cancer Research,* 73**,** 3097-3108.

Wang, T., Arifoglu, P., Ronai, Z. e. & Tew, K. 2001. Glutathione s-transferase p1-1 (gstp1-1) inhibits c-jun n-terminal kinase (jnk1) signaling through interaction with the c terminus. *The Journal of biological chemistry,* 276**,** 20999-1003.

Wang, X.-J., Sun, Z., Villeneuve, N. F., Zhang, S., Zhao, F., Li, Y., Chen, W., Yi, X., Zheng, W., Wondrak, G. T., Wong, P. K. & Zhang, D. D. 2008a. Nrf2 enhances resistance of cancer cells to chemotherapeutic drugs, the dark side of nrf2. *Carcinogenesis,* 29**,** 1235-1243.

Wang, Z., Smith, K. S., Murphy, M., Piloto, O., Somervaille, T. C. P. & Cleary, M. L. 2008b. Glycogen synthase kinase 3 in mll leukaemia maintenance and targeted therapy. *Nature,* 455**,** 1205.

Wasserman, W. W. & Fahl, W. E. 1997. Functional antioxidant responsive elements. *Proceedings of the National Academy of Sciences,* 94**,** 5361-5366.

Westcott, P. & To, M. 2012. The genetics and biology of kras in lung cancer. *Chinese journal of cancer,* 32**,** 63-70.

Wimuttisuk, W., West, M., Davidge, B., Yu, K., Salomon, A. & Singer, J. D. 2014. Novel cul3 binding proteins function to remodel e3 ligase complexes. *BMC Cell Biology,* 15**,** 28.

Winterbourn, C. C. 2013. Chapter one - the biological chemistry of hydrogen peroxide. *In:* CADENAS, E. & PACKER, L. (eds.) *Methods in enzymology.* Academic Press.

Woods, Y. L., Cohen, P., Becker, W., Jakes, R., Goedert, M., Wang, X. & Proud, C. G. 2001. The kinase dyrk phosphorylates protein-synthesis initiation factor eif2bepsilon at ser539 and the microtubule-associated protein tau at thr212: Potential role for dyrk as a glycogen synthase kinase 3-priming kinase. *Biochemical Journal,* 355**,** 609-615.

Wu, G., Xu, G., Schulman, B. A., Jeffrey, P. D., Harper, J. W. & Pavletich, N. P. 2003. Structure of a β-trcp1-skp1-β-catenin complex: Destruction motif binding and lysine specificity of the scfβ-trcp1 ubiquitin ligase. *Molecular Cell,* 11**,** 1445-1456.

Wu, T., Zhao, F., Gao, B., Tan, C., Yagishita, N., Nakajima, T., K Wong, P., Chapman, E., Fang, D. & D Zhang, D. 2014. Hrd1 suppresses nrf2-mediated cellular protection during liver cirrhosis. *Genes & development,* 28**,** 708-22.

Wu, X. & Rapoport, T. A. 2018. Mechanistic insights into er-associated protein degradation. *Current Opinion in Cell Biology,* 53**,** 22-28.

Xia, C., Meng, Q., Liu, L.-Z., Rojanasakul, Y., Wang, X.-R. & Jiang, B.-H. 2007. Reactive oxygen species regulate angiogenesis and tumor growth through vascular endothelial growth factor. *Cancer Research,* 67**,** 10823-10830.

Yamamoto, K., Ichijo, H. & Korsmeyer, S. J. 1999. Bcl-2 is phosphorylated and inactivated by an ask1/jun n-terminal protein kinase pathway normally activated at g(2)/m. *Molecular and Cellular Biology,* 19**,** 8469-8478.

Yeager, R. L., Reisman, S. A., Aleksunes, L. M. & Klaassen, C. D. 2009. Introducing the "tcdd-inducible ahr-nrf2 gene battery". *Toxicological Sciences,* 111**,** 238-246.

Yoshida, T., Ohkumo, T., Ishibashi, S. & Yasuda, K. 2005. The 5′-at-rich half-site of maf recognition element: A functional target for bzip transcription factor maf. *Nucleic Acids Research,* 33**,** 3465-3478.

You, A., Nam, C.-w., Wakabayashi, N., Yamamoto, M., Kensler, T. W. & Kwak, M.-K. 2011. Transcription factor nrf2 maintains the basal expression of mdm2: An implication of the regulation of p53 signaling by nrf2. *Archives of Biochemistry and Biophysics,* 507**,** 356-364.

Zeeshan, H. M. A., Lee, G. H., Kim, H.-R. & Chae, H.-J. 2016. Endoplasmic reticulum stress and associated ros. *International Journal of Molecular Sciences,* 17**,** 327.

Zhang, C., Wang, H.-J., Bao, Q.-C., Wang, L., Guo, T.-K., Chen, W.-L., Xu, L.-L., Zhou, H.-S., Bian, J.-L., Yang, Y.-R., Sun, H.-P., Xu, X.-L. & You, Q.-D. 2016. Nrf2 promotes breast cancer cell proliferation and metastasis by increasing rhoa/rock pathway signal transduction. *Oncotarget,* 7**,** 73593-73606.

Zhang, D., Lo, S.-C., Sun, Z., M Habib, G., W Lieberman, M. & Hannink, M. 2005. Ubiquitination of keap1, a btb-kelch substrate adaptor protein for cul3, targets keap1 for degradation by a proteasome-independent pathway. *Journal of Biological Chemistry,* 280**,** 30091-9.

Zhang, D. D. & Hannink, M. 2003. Distinct cysteine residues in keap1 are required for keap1-dependent ubiquitination of nrf2 and for stabilization of nrf2 by chemopreventive agents and oxidative stress. *Molecular and Cellular Biology,* 23**,** 8137-8151.

Zhang, D. D., Lo, S.-C., Cross, J. V., Templeton, D. J. & Hannink, M. 2004. Keap1 is a redox-regulated substrate adaptor protein for a cul3-dependent ubiquitin ligase complex. *Molecular and Cellular Biology,* 24**,** 10941-10953.

Zhang, H., Davies, K. J. A. & Forman, H. J. 2015. Oxidative stress response and nrf2 signaling in aging. *Free radical biology & medicine,* 88**,** 314-336.

Zhang, H., Zhao, W., Venkataraman, S., E C Robbins, M., Buettner, G., C Kregel, K. & W Oberley, L. 2002. Activation of matrix metalloproteinase-2 by overexpression of manganese superoxide dismutase in human breast cancer mcf-7 cells involves reactive oxygen species. *Journal of Biological Chemistry,* 277**,** 20919-26.

Zhang, J., Hosoya, T., Maruyama, A., Nishikawa, K., Maher, Jonathan M., Ohta, T., Motohashi, H., Fukamizu, A., Shibahara, S., Itoh, K. & Yamamoto, M. 2007a. Nrf2 neh5 domain is differentially utilized in the transactivation of cytoprotective genes. *The Biochemical Journal,* 404**,** 459-466.

Zhang, J., Ohta, T., Maruyama, A., Hosoya, T., Nishikawa, K., Maher, J. M., Shibahara, S., Itoh, K. & Yamamoto, M. 2006a. Brg1 interacts with nrf2 to selectively mediate ho-1 induction in response to oxidative stress. *Molecular and Cellular Biology,* 26**,** 7942-7952.

Zhang, Y., Crouch, Dorothy H., Yamamoto, M. & Hayes, John D. 2006b. Negative regulation of the nrf1 transcription factor by its n-terminal domain is independent of keap1: Nrf1, but not nrf2, is targeted to the endoplasmic reticulum. *Biochemical Journal,* 399**,** 373-385.

Zhang, Y., Kensler, T. W., Cho, C. G., Posner, G. H. & Talalay, P. 1994. Anticarcinogenic activities of sulforaphane and structurally related synthetic norbornyl isothiocyanates. *Proceedings of the National Academy of Sciences,* 91**,** 3147-3150.

Zhang, Y., Kobayashi, A., Yamamoto, M. & Hayes, J. 2009. The nrf3 transcription factor is a membrane-bound glycoprotein targeted to the endoplasmic reticulum through its n-terminal homology box 1 sequence. *Journal of Biological Chemistry,* 284**,** 3195-210.

Zhang, Y., Lucocq, John M., Yamamoto, M. & Hayes, John D. 2007b. The nhb1 (n-terminal homology box 1) sequence in transcription factor nrf1 is required to anchor it to the endoplasmic reticulum and also to enable its asparagine-glycosylation. *The Biochemical Journal,* 408**,** 161-172.

Zhang, Z., Wang, S., Zhou, S., Yan, X., Wang, Y., Chen, J., Mellen, N., Kong, M., Gu, J., Tan, Y., Zheng, Y. & Cai, L. 2014. Sulforaphane prevents the development of cardiomyopathy in type 2 diabetic mice probably by reversing oxidative stress-induced inhibition of lkb1/ampk pathway. *Journal of Molecular and Cellular Cardiology,* 77**,** 42-52.

Zhao, Z., Chen, Y., Wang, J., Sternberg, P., Freeman, M. L., Grossniklaus, H. E. & Cai, J. 2011. Age-related retinopathy in nrf2-deficient mice. *PLOS ONE,* 6**,** 19456.

Zhou, L., Zhang, H., Davies, K. J. A. & Forman, H. J. 2018. Aging-related decline in the induction of nrf2-regulated antioxidant genes in human bronchial epithelial cells. *Redox Biology,* 14**,** 35-40.

Zhou, S., Ye, W., Zhang, M. & liang, J. 2012. The effects of nrf2 on tumor angiogenesis: A review of the possible mechanisms of action. 22**,** 149-160.

Zhu, Q.-G., Zhang, S.-M., Ding, X.-X., He, B. & Zhang, H.-Q. 2017. Driver genes in non-small cell lung cancer: Characteristics, detection methods, and targeted therapies. *Oncotarget,* 8**,** 57680-57692.

Zipper, L., M. & Mulcahy, T., R. 2002. The keap1 btb/poz dimerization function is required to sequester nrf2 in cytoplasm. *Journal of Biological Chemistry,* 277**,** 36544-52.

Zu, X., Yan, R., Ma, J., Liao, D.-F. & Cao, D. 2009. Akr1b10: A potential target for cancer therapy. *Bioscience Hypotheses,* 2**,** 31-33.

8.0 Appendix

8.1 Appendix I

Mutation	Primer Sequence (5'-3')
S335A	GAA TTC AAT GAC TCT GAC **GCT** GGC ATT TCA CTG AAC ACG
S338A	GAC TCT GAC TCT GGC ATT **GCG** CTG AAC ACG AGT CCC AGT CCC
S342A	GGC ATT TCA CTG AAC ACG **GCT** CCC AGC CGA GCG TCC CCA
S347A	ACG AGT CCC AGC CGA GCG **GCC** CCA GAG CAC TCC GTG GAG
V345R	CTA AAC ACA AGT CCC AGT **CGG** GCA TCA CCA GAA CAC TCA

Table 8.1.1: Oligonucleotide primers sequences for the generation of site mutant plasmids

8.2 Appendix II

Cell line name	Guide sequence (5'-3')
NRF2 GOF H1299 / NRF2 GOF H1299	TGG AGG CAA GAT ATA GAT CT
KEAP1 KO H1299	CAC CGC ACC TTC AGC TAC ACC CTG G

Table 8.2.1: Guide sequences for the generation of CRISPR/Cas9 cell lines

8.3 Appendix III

Antibody	Company and catalogue number	Dilution
Rabbit anti-NRF2	Abcam (ab62352)	1:1000
Rabbit anti-KRAS	Abcam (ab180772)	1:750
Rabbit anti-NQO1	In house	1:2000
Rabbit anti-Phospho GSK-3α/β (Ser9/21)	Cell signalling Technology (93318)	1:1000
Rabbit anti-Phospho AKT (Thr 308)	Cell signalling Technology (4056S)	1:1000
Rabbit anti-GSK-3α/β	Cell signalling Technology (5676S)	1:1000
Rabbit anti-Phospho AKT (Ser473)	Cell signalling Technology (4060S)	1:1000
Rabbit anti-G6PD	Cell signalling Technology (12263S)	1:1000
Rabbit anti-PGD	Cell signalling Technology (13389S)	1:1000
Rabbit anti-GAPDH	Cell signalling Technology (2118S)	1:2000
Rabbit anti-AKT (Pan)	Cell signalling Technology (4691S)	1:1000
Rabbit anti-AKR1B10	Home made	1:2000
Rabbit anti-AKR1C1	Home made	1:2000
Rabbit anti-GCLM	ThermoFisher Scientific (PA5-26111)	1:1000
Rabbit anti-HMOX-1	BioVison (3391-100)	1:1000
Rabbit anti-ME1	Invitrogen (PA5-21550)	1:1000
Rabbit anti- GPX2	Abcam(ab137431)	1:1000
Rabbit anti-SLC7A11	Abcam (ab37185)	1:1000
Rabbit anti- GCLC	Invitrogen (PA5-16581)	1:1000
Rat anti- KEAP1 Clone 144	Merck (MABS514)	1:1000
Mouse anti-Lamin A/C	Sigma-Aldrich (SAB4200236)	1:1000
Mouse anti-Lamin B2	Santa Cruz Biotechnology (sc-377379)	1:1000
Mouse anti-P62	Merck (MABC32)	1:1000
Rabbit anti-PGAM5	Merck (ABC517)	1:1000
Rabbit anti-TSC22D4	ThermoFisher Scientific (PA5-40333)	1:1000

Table 8.3.1: Source and dilution of all primary antibodies

8.4 Appendix IV

Gene	Catalogue number
AKR1C2	Hs04194036_gH
NQO1	Hs02512143_s1
KEAP1	Hs00202227_m1
NFE2L2	Hs00975961_g1
GCLC	Hs00155249_m1
GCLM	Hs00157694_m1
GPX2	Hs01591589_m1
SLC7A11	Hs00921938_m1
G6PD	Hs00166169_m1
PGD	Hs00427230_m1
ME1	Hs00159110_m1
IDH1	Hs00271858_m1
HMOX1	Hs01110250_m1
GAPDH	Hs02758991_g1
AKR1B10	Hs00252524_m1
AKR1C1	Hs04230636_Sh
AKR1C3	Hs00366267_m1
TALDO1	Hs00997203_m1
KRAS	Hs00364284_g1
18S rRNA	Hs99999901_s1

Table 8.4.1: List and catalogue number of TaqMan probes sets purchased from Applied Biosystems

Gene	TaqMan primer probe sequences (5'-3')
ACTIN FW primer ACTIN RV primer ACTIN probe	GCGCGGCTACAGCTTCA TCTCCTTAATGTCACGCA CACCACGGCCGAGCGGG
HMOX1 FW primer HMOX1 RV primer HMOX1 probe	AGTGCCACCAAGTTCAAGC TATCACCCTCTGCCTGACTG ACCGCTCCCGCATGAACTCC

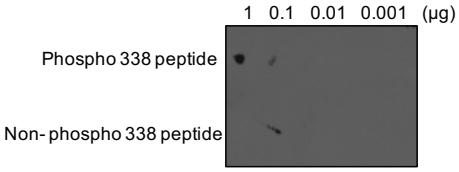
Table 8.4.2: List of sequences of customised TaqMan primers and probes

8.5 Appendix V

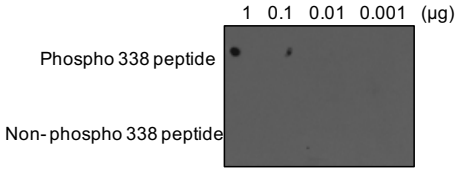
Amino acid	Abbreviation	Letter
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamic acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Try	W
Tyrosine	Tyr	Y
Valine	Val	V

Table 8.5.1: List of amino acid names, abbreviation and letter codes

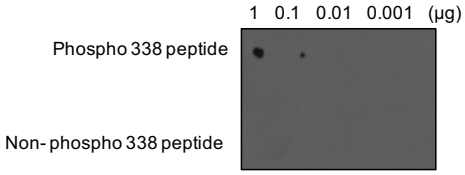
8.6 Appendix VI



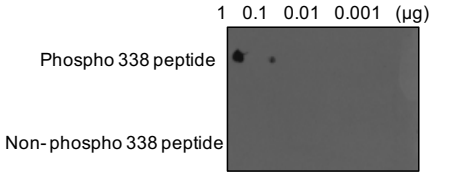
1:500 (3rd bleed) P-S342 Nrf2



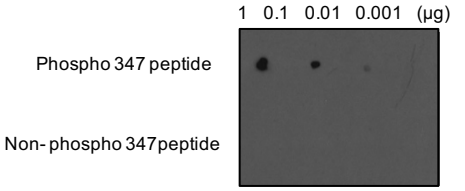
1:1000 (3rd bleed) P-S342 Nrf2



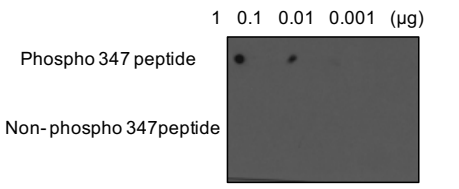
1:500 (3rd bleed) P-S347 Nrf2



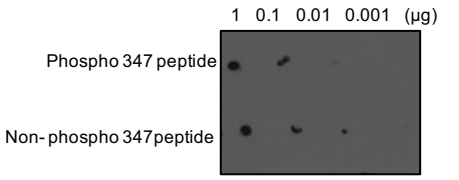
1:1000 (3rd bleed) P-S347 Nrf2



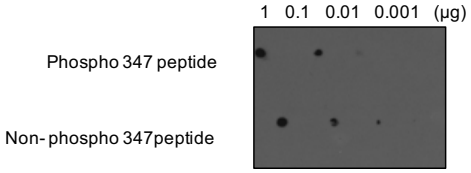
1:500 (3rd bleed) P-S342 Nrf2



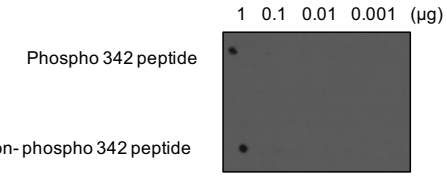
1:1000 (3rd bleed) P-S342 Nrf2



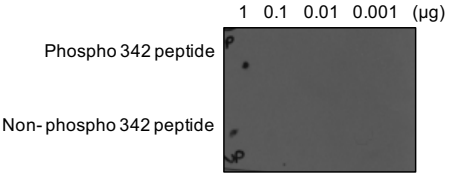
1:500 (3rd bleed) P-S347 Nrf2



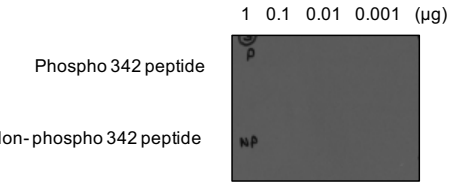
1:1000 (3rd bleed) P-S347 Nrf2



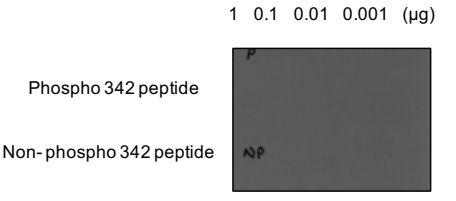
1:500 (3rd bleed) P-S342 Nrf2



1:1000 (3rd bleed) P-S342 Nrf2



1:500 (3rd bleed) P-S347 Nrf2



1:1000 (3rd bleed) P-S347 Nrf2

Figure 8.6.1: Dot blot optimization for the in house phospho-specific antibodies

Dilutions of the peptides are indicated above the images; the name of the peptide is given down the left-hand side of the image and the antibodies (dilution) used to probe the membrane is indicated below the image. Dilutions of phospho peptides and non-phospho peptides were dotted on to nitrocellulose membranes. Membranes were blocked in 5% BSA for 1 hr and then incubated in primary antibody for 2 hr. Membranes were then incubated in 1:1000 sheep HRP-conjugated secondary antibody followed by visualization with ECL. These images are representative of three biological replicates.

8.7 Appendix VII

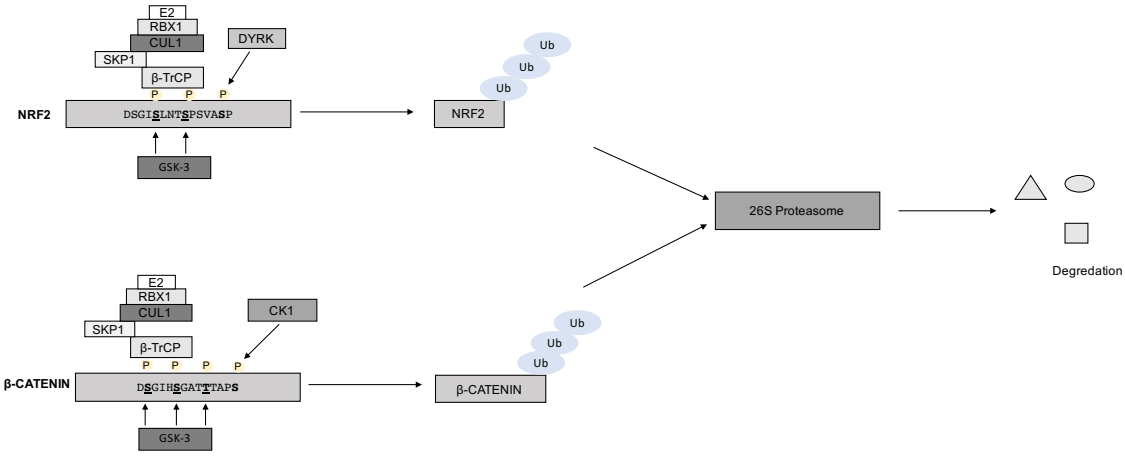


Figure 8.7.1: β-TrCP mediated degradation of NRF2 is similar to that of β-Catenin, which also requires priming of GSK-3

Schematic highlighting the similarities in the β-TrCP mediated degradation of NRF2 and that of β-catenin. Priming by DYRK leads to enhanced phosphorylation of NRF2 by GSK-3 leading to the formation of the phosphodegron to which the substrate adaptor protein β-TrCP binds. Once β-TrCP is bound to NRF2 is recruits the CUL1 ubiquitin machinery complex that labels NRF2 with ubiquitin for subsequent degradation through the 26S proteasome route. This pathway of degradation for NRF2 is extremely similar to the well characterised β-catenin degradation pathway. Priming by CK1 leads to enhanced phosphorylation of β-catenin by GSK-3 leading to the formation of the phosphodegron to which β-TrCP binds. Once bound β-TrCP recruits the CUL1 ubiquitin machinery complex leading to degradation of β-catenin through the 26S proteasome. Adapted from Robertson et al. (2018).

8.8 Appendix VIII

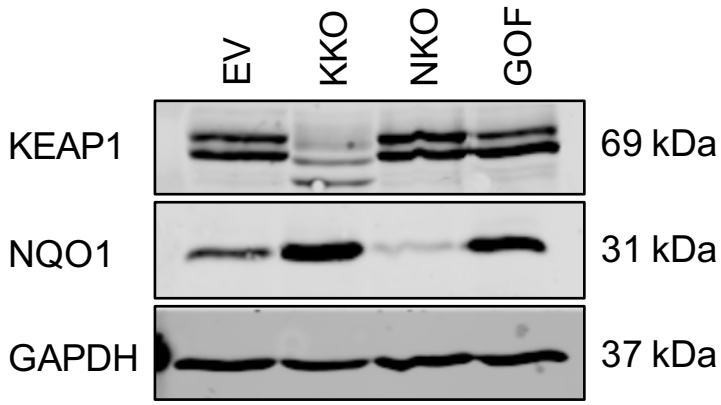
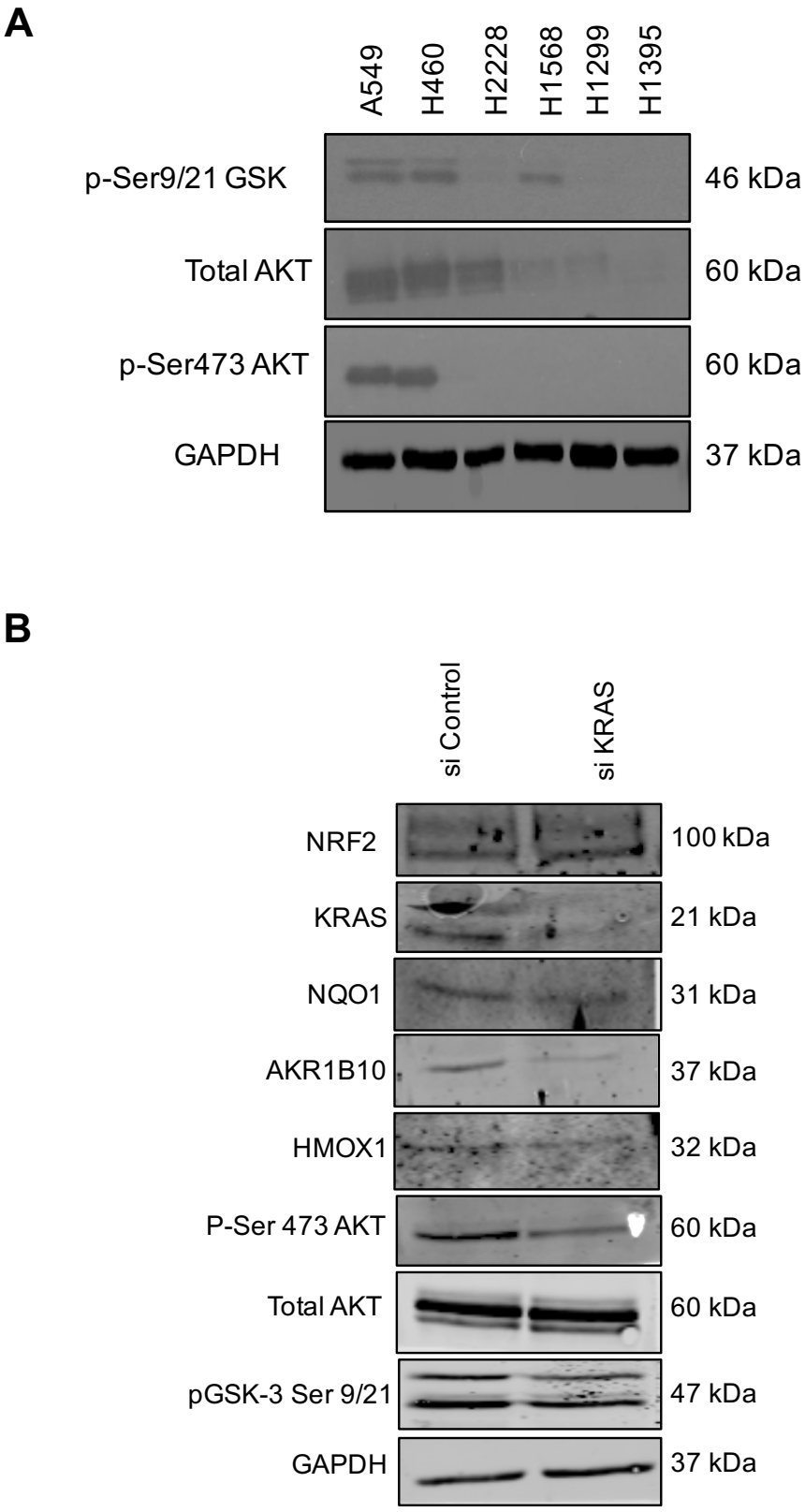


Figure 8.8.1: KKO H1299 cells express lower protein levels of KEAP1

Protein analysis of KEAP1 and NQO1 across the H1299 CRISPR panel of cell lines. This image is taken from one experiment and representative of three independent biological replicates.

8.9 Appendix IX



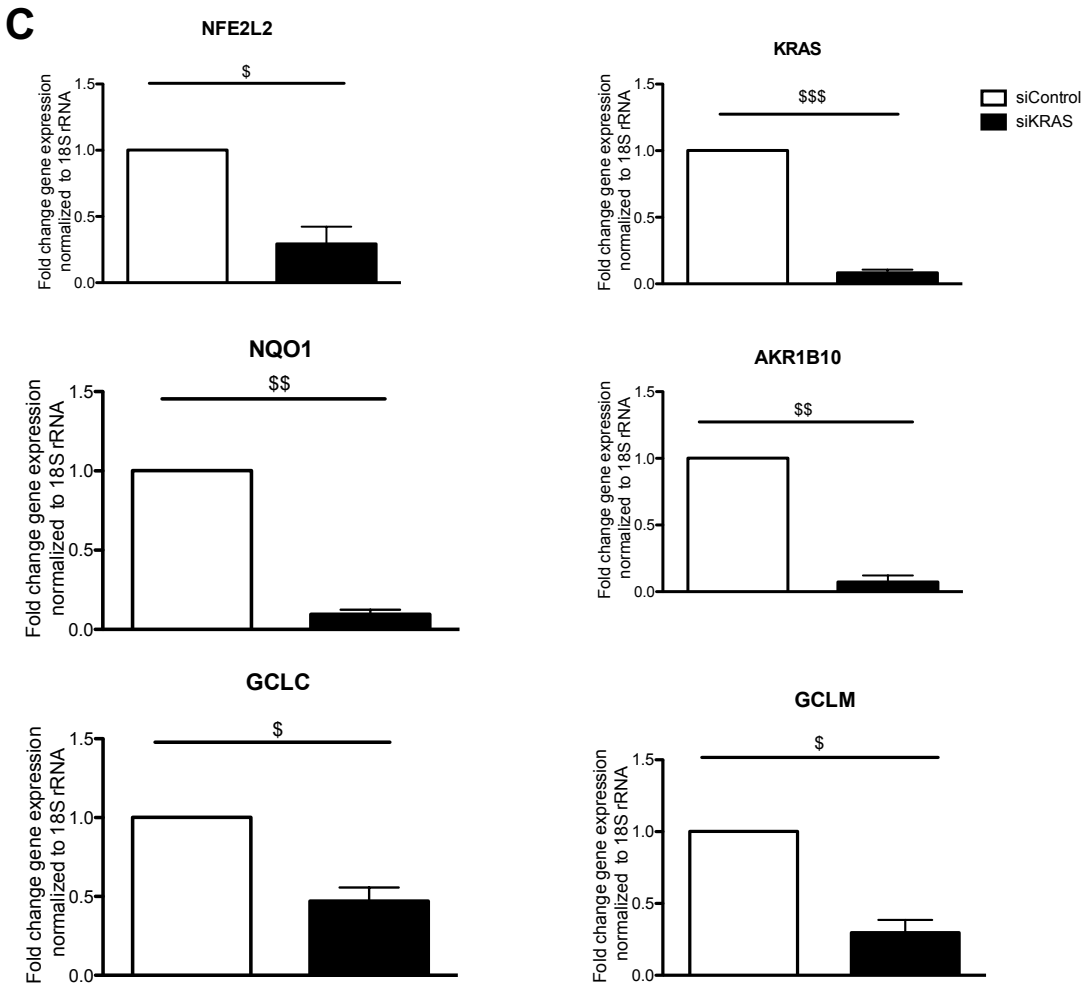

C
NFE2L2
$
KRAS
$$$
siControl
siKRAS
NQO1
$$
AKR1B10
$$
GCLC
$
GCLM
$
Fold change gene expression normalized to 18S rRNA
1.5
1.0
0.5
0.0

Figure 8.9.1: AKT signalling is increase in cell lines harbouring a KRAS mutation and knock down of KRAS leads to decreased NRF2-target gene expression.

*(A) Protein analysis of the AKT signalling pathway in: A549, H460, H2228, H1568, H1299 and H1395 cells. Cell lines are indicated across the top of the image, the names of the proteins being analysed are given down the left-hand side of the image and their corresponding molecular weights on the right-hand side. This image is one representative image of three independent biological replicates. (B) Protein analysis of NRF2-target genes and AKT signalling in DLD1 cells upon siRNA knock down of KRAS. The results for the siRNA control are given in the left lane of the western blot and the siRNA KRAS knockdown results are given in the right line of the image. The names of the proteins being analysed are given down the left-hand side of the image and their corresponding molecular weights on the right-hand side. This image is a representative of three independent biological replicates. (C) mRNA analysis of NRF2-target genes after siRNA knockdown of KRAS in DLD1 cells. The white bar corresponds to the siRNA control cells and the black bar represents the siRNA KRAS knockdown cells. The values displayed on the graphs are the mean value from three replicate experiments and the associated SEM. All statistical analysis was carried out using GraphPad Prism 5 software. All significant increases are denoted as follows: *, P< 0.05; **, P<0.01; *** P< 0.001. P values > 0.05 were deemed not significant and are denoted by "ns". Significant decreases are denoted with a $ sign.*

8.10 Appendix X

Cell line name	Name of the gene disrupted	Sequencing results	Predicated affect
Clone 8 KKO	KEAP1	Homozygous in-frame deletion 87bp (M1)	KEAP1 protein not produced
Clone 40 NKO	NFE2L2	Heterozygous NRF2 KO deletion 20bp (L23) and 7bp (D29)	NRF2 protein not produced
Clone 31 GOF	NFE2L2	Heterozygous in-frame deletion 3bp (D29) and 24bp (R25-V32)	KEAP1 binding to NRF2 impaired and NRF2 levels increased

Table 8.10.1: Sequencing results for the CRISPR/Cas9 H1299 cells

Sequencing was kindly carried out by Mr Thomas McCartney. The name of the cell lines are given in the first left-hand side of the column, followed by the name of the gene which has been altered, the sequencing results and the predicated effect of the disruption; in the adjacent columns.